DOCUMENT RESUME

ED 326 215 IR 052 714

TITLE IFLA General Conference 1988. Division of Bibliographic Control. Section on Cataloguing; Section on Classification and Indexing; Section on Bibliography.

INSTITUTION International Federation of Library Associations, The Hague (Netherlands).

PUB DATE Sep 88

NOTE 18p.; Papers presented at the International Federation of Library Associations (IFLA) General Conference (54th, Sydney, Australia, August 30-September 3, 1988).

PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Bibliographies; *Cataloging; *Classification; Cultural Influences; Databases; Foreign Countries; *Indexing; Librarians; Library Automation; *Library Catalogs; Machine Readable Cataloging; National Libraries; Online Catalogs; Research Libraries; Technological Advancement

IDENTIFIERS *Dewey Decimal Classification; *National Bibliographies; Retrospective Conversion (Library Catalogs); Universal Decimal Classification

ABSTRACT

The nine papers in this compilation focus on cataloging, classification, and indexing: (1) "Living Together in the Australian Bibliographic Network: A Member's View" (Tom Cochrane, Australia); (2) "PICA (Project for Integrated Catalogue Automation), from Project to Institution: A Participant's View" (Jacques van Gent & Cees Klapwijk, The Netherlands); (3) "Classification Online: Present and Future" (Nancy Williamson, Canada); (4) "Australian Directions in Subject Access" (Elaine Hall, Australia); (5) "Retrospective Conversion in National and Research Libraries: The Australian Experience" (Peter Haddad, Australia); (6) "Retrospective Conversion in the Libraries of Socialist Countries: A Survey" (N. P. Igumnova, USSR); (7) "Problems of Retrospective Conversion in National and Research Libraries: The Experience of the Western Countries--The Conversion of the British Library Catalogue to Machine Readable Form" (Pat Oddy, United Kingdom); (8) "Problems of Retrospective Conversion in National and Research Libraries: The Experience of the Western Countries--Integrating the Catalogue General des Livres Imprimes of the Bibliotheque Nationale into the BN-OPALE Data Base" (Annick Bernard, France); and (9) "Problems of Retrospespective Conversion in National and Research Libraries: The Experience of the Western Countries--The Spanish Experience" (Xavier Agenjo, Spain). (CGD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED326215

U.S. DEPARTMENT OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

54th IFLA
GENERAL CONFERENCE
SYDNEY, AUSTRALIA
30 AUGUST - 3 SEPTEMBER 1988

Division of Bibliographic Control

SECTION — CATALOGUING

Session: 132 Friday 2 September 0900 - 1130

Living together in the Australian Bibliographic Network: a member's view *(31-CAT-1-E)*

Tom Cochrane, Chief Librarian, Queensland Institute of Technology, Brisbane, Queensland 4000, Australia

PICA, from project to institution. A participant's view *(32-CAT-2-E)*

Jacques J.M. van Gent and Cees A. Klapwijk, State University Library, Leiden, The Netherlands

SECTION — CLASSIFICATION AND INDEXING

Session: 120 Thursday 1 September 1500 - 1730

Classification Online: present and future *(21-CLASS-1-E)*

Nancy J. Williamson, Faculty of Library and Info. Science, University of Toronto, Canada

Australian directions in subject access *(22-CLASS-2-E)*

Elaine Hall, State Library of New South Wales, Sydney, Australia

SECTION — BIBLIOGRAPHY

Session: 103 Thursday 1 September 0900 - 1130

Retrospective conversion in national and research libraries: the Australian experience *(5-BIBL-1-E)*

Peter Haddad, Chief Librarian, Control, National Library of Australia, Canberra, ACT 2600

Retrospective conversion in the libraries of Socialist countries: a survey *(18-BIBL-2-E)*

N.P. Igumnova, Director, All-Union State Library of Foreign Literature, Moscow, USSR

Problems of retrospective conversion in national and research libraries: the experience of the Western countries.

— The conversion of the British Library catalogue to machine readable form *(43-BIBL-3a-E)*

Pat Oddy, Head, BLC Conversion Project, British Library Bibliographic Services, London, UK

— Integrating the Catalogue général des livres imprimés 1970–1979 of the Bibliothèque Nationale into the BN-OPALE data base *(43-BIBL-3b-E)*

Annick Bernard, Chef de Projet, Bibliothèque Nationale, Paris

— The Spanish experience *(43-BIBL-3c-E)*

Xavier Agenjo, Chief of the Service of the Union Catalog of Spanish Heritage, National Library, Madrid

IR052714

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
A.L. Van Wesemael

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Data conversion by Saztec Information Services as a contribution to IFLA/LAA 1988

LIVING TOGETHER IN THE AUSTRALIAN BIBLIOGRAPHIC NETWORK: A MEMBER'S VIEW

Tom Cochrane, *Chief Librarian, Queensland Institute of Technology, Brisbane, Queensland 4000, Australia*

ABSTRACT

This paper briefly outlines the development of the Australian Bibliographic Network and the experience of one member library's participation in that network from the early 1980s until now. Reasons for joining the network in the first place are given, and the author assesses the benefits and drawbacks of this membership. After describing these, a judgement is made about the benefit of participation, and the institution's stake in the future of the network.

INTRODUCTORY REMARK

For obvious reasons I have attempted to minimise the use of initialisms and acronyms, a use at which librarians everywhere seem to excel. However there are three which unavoidably recur throughout my paper, with which the reader or listener should be familiar. These are:

ABN — the Australian Bibliographic Network
QIT — Queensland Institute of Technology
MARC — Machine Readable Cataloguing

1. Introduction

Australia is a federation of six states and two territories, comprising 16 million people. Of these 16 million, 60% live in the two states of New South Wales and Victoria, and 62% live within all state and national capitals. As most visitors soon realise, it is a country where distances are great, and the population is small. By way of comparison, Australia has a population density per square kilometre of 2, compared to Canada's 3, the United States' 25, Japan's 318, and the Netherlands' 353.[1]

The country's libraries fall into familiar categories: national library, six state libraries, academic libraries in Australia's universities and colleges (approximately 80 in the higher education sector), public libraries, school libraries and special purpose libraries.

The 1960s and 70s saw the National Library initiate national services for library cataloguing, beginning with a card service inaugurated in 1967. The idea of a national network for shared cataloguing received encouragement from the Australian library community during the 1970s, and the Australian government eventually approved the establishment of the Network in August 1981. It is important to note that it began as a National Library activity (the government approved and funded it only on this basis), and has always been expected to recover its operating costs.

The basis for the system is the computer software developed for the purpose of shared cataloguing in the U.S. state of Washington and originally known as the Washington Library Network (WLN). Features included relative sophistication of authority control, and relative intolerance of duplicate (or multiple) bibliographic records, in contrast to some other utilities providing similar services.

Starting with six full participants (i.e. libraries who add records and location statements to the data base in addition to using it to provide cataloguing products and information), the network added a further 20 in the first year, a further 25 in the second, and by mid 1984, had 72 full participants and 105 dial up customers. By early 1988, after six years of operation, these numbers had become 137 and 572.[2]

2. Queensland Institute of Technology

One of these participants is the Queensland Institute of Technology, which joined the network in mid 1983. Queensland Institute of Technology (or QIT) is the only Institute of Technology in Queensland, which is Australia's third largest state in population terms, and second largest state in area.

The Institute, which offers 60 courses, mostly at degree and postgraduate level, has 10,000 students and 1,000 staff, and teaches Engineering, Science, Health Science, Information Technology, Business, Built Environment (Architecture, Building and Planning), and Law. It is the only higher education campus in the central district of Brisbane, which is Australia's third largest city, with a population of 1.2 million.

The institute, with its Library, is just over twenty years old, and is based on an old technical college with no previous library provision of any size. Like many higher education libraries in Australia its relative recency has meant a generally modest collection size.

One of the advantages associated with this feature (a limitation in other ways, of course) is the fact that half of the present collection has been catalogued after the library moved to machine readable cataloguing at the end of the 1970s; and the remainder was retrospectively converted by a bureau service, so that all holdings were converted by the end of 1982. Before QIT commenced negotiation with ABN therefore, it was in the fortunate position of having records for its monograph, non-book, and serial collections all in machine readable form, although the records for serials were separately filed in a non-MARC format (which was therefore to ABN, not immediately acceptable for retrospective conversion).

3. When and How QIT Joined ABN

In the first half of 1983, QIT arranged for a complete pass of its bureau-held machine readable file to extract all monograph records, and these, something over 100,000, were sent to ABN at the National Library in mid year. Non-book records were excluded at this stage, because ABN could not handle certain formats in this mode. Similarly, QIT's non-MARC serials records, held at QIT's own computer centre, were not part of the original conversion.

All things considered, the conversion of monographs and transition to a new online cataloguing operation went reasonably well at QIT when it occurred in July 1983. Similarly, the conversion of non book records, which ultimately took place in November 1986, was relatively trouble free. Not so the conversion of the non-MARC serials file, more than half of which was rejected for review from the initial matching process, a review which has taken almost two years since the first pass was made in August 1986. A considerable labour-cost was associated with this review, but this was an unavoidable consequence of an earlier decision (1978) by QIT to depart from MARC standards for its serials file.

It is worth nothing that QIT has never been completely dependent upon ABN for its "end user" catalogue products. From July 1984, an online catalogue was made available locally, (a by-product of the QIT s Dataphase circulation control system), and ABN microfiche listings were available as back-ups to this principal, (and more popular) tool. The online catalogue was, and is, of course, maintained by ABN tape updates.

Following the retrospective conversions for the various media, regular procedures for the ongoing cataloguing of items were instituted as follows:

— at the time of placing an order, the network is searched. If a full or acceptable level of record is located, a "QIT on order" statement is added to the holdings field. Henceforth, any user of the network can detect that that item is "on order" for QIT. If no record exists, an interim record is established, to which the QIT "on order" statement is then added;
— when the work is in hand, the "on order" statement is replaced by a call number, and the record upgraded if this is still required;
— regular products received include tapes of new holdings and bibliographic changes, microfiche catalogues by author, title, and subject (and a main "register") and printouts of new headings requiring verification for the purpose of authority control.

4. Why QIT joined ABN

There were several reasons why QIT joined the national network in 1983:

First, we strongly believed that it would improve the standard of our own cataloguing and catalogue products.

Second, we believed, though perhaps with less certain conviction, that it would be cost beneficial compared to our existing machine readable cataloguing procedures.

Third, we were committed, practically and professionally, to the idea of the national network, and looked forward to being a part of it.

Fourth, it has relatively easy for QIT to do. We had neither the problem of a large unconverted file, nor the problem of an alternative and consuming investment in a local product. We were already paying for batch machine readable cataloguing services, — it was therefore, simply a matter of transferring payment to ABN.

5. The Benefits

Well, what, looking back over the last five years, have been the benefits to QIT of this participation, this "living together" with other libraries in ABN?

First and foremost, our best expectations regarding visible costs of the service we get from ABN have been realised. In 1983, bureau services represented 3% of total library budget. By 1988, this has

fallen to 1.33%, an effective halving of the costs of automated cataloguing. Admittedly, against this must be set against the costs, in staff time, sometimes high, of each retrospective conversion, but these were bound to occur particularly with any project to bring our serials records up to MARC standard, whether ABN was involved or not. Soon after converting to ABN, one full time cataloguing professional was removed from the Cataloguing section. However, in 1987, this position was effectively restored, partly because ongoing authority work generated by the network was not being adequately done. Thus, overall staffing levels have been more or less stable, bearing in mind that QIT previously used a bureau service. However, the staffing profile has seen a move towards paraprofessional and away from professional labour, so costs as a proportion of library salaries have marginally declined.

Second, our hopes of achieving improving "hit rates" for centralised copy have also been realised. With our previous system, 57% of items searched provided a record of sufficient quality to use. This figure rose to 75% when we first joined the network, and has since risen to 82%. It is conceivable that the figure will improve still further.

Third, the professionalism of our own cataloguing operation has been improved by joining this online community of peers. Dialogues with the bibliographic review staff of ABN are both instructive and stimulating, conducted, as they have been from the start, in a friendly and informal way.

Fourth, the burden of maintaining the copious documentation that goes with cataloguing procedures is a good deal lighter for members of the network, with the provision of a generally high standard of the requisite manuals for subject cataloguing and so on.

Fifth, part of living in this network means a consistency in authority work. QIT gains the advantage of other users' authority verifications and vice versa. (There is, however, a specific problem involved in authority work following retrospective conversions, which I refer to subsequently).

Sixth, with the exception of the fire which disabled ABN, and many other parts of the National Library in 1985, the system itself, has, generally speaking been reliable during QIT's use of it. Apart from the fire, our (totally dependent) cataloguing work has never been seriously disrupted by hardware or communication problems.

Seventh, technical and bibliographic back up through the daily telephone advice available from ABN has been of a high order.

Eighth, though hard to quantify, it has been rewarding in the most general terms to be part of the network as it grows. QIT now shares common concerns, on which it can exchange views, with libraries all over the continent.

Finally, the system itself offers the immediate and future capability of providing services other than shared cataloguing. From the beginning, we have used it for inquiry and interlibrary loan purposes, and were (and remain) interested in its acquisitions capability. We shall not hesitate to avail ourselves of the interlibrary lending module when it becomes available.

6. The Drawbacks

So much for the benefits. What have the drawbacks associated with our cohabitation with other libraries been?

I have just mentioned interlibrary loan, and it is in this area that we have experienced a major difficulty. Soon after our retrospective conversion of monograph records we began to experience an extraordinary increase in interlibrary loan requests for books. This was of the order of 100% over the first two years of our membership.

We had anticipated some increased demand, but nothing of this order. It was just as well that we had not also loaded our serials records, or the demand would have been higher. This problem was perhaps exacerbated by the fact that QIT was, and remains the only ABN participant of any size from Australia's third largest state, and a regional load has fallen to us perhaps more than it may have. It must be said though, that this demand, which has had the effect of making QIT a "net lender" (a library that lends more than it borrows in the interlibrary network) has not worked too much to our disadvantage, as the Australian library community has moved slowly towards a greater acceptance of the principle of sharing the real costs of the interlending system.

This particular drawback had nothing to do with cataloguing per se. More important here, have been:

first, the unmanageable burden of retrospective authority work which arose from our initial addition of records to the network. This was so large that we simply never tackled it, like many other members of the network;

second, it has inconvenient to us that ABN lacked the technical capacity to accept non-book MARC data from our previous system, so that we in fact waited for 3 1/2 years before these could be added;

third, living in this network has, like most network arrangements, meant that we have had, on occasion, to subjugate individual library requirements to network constraints and technical difficulties. We have found, for example, that we want to treat records in a particular way for QIT purposes, but are constrained by the network record. This is a direct outcome, of course, of the approach of having the one unique record for all users. We have also had to accept lengthy delays in previously agreed schedules for retrospective conversions, and in the implementation of needed system improvements;

fourth, some of the technical aspects of the system have been clumsy, and when these also seem relatively simple to fix, it has been irritating that they take a long time to be resolved because of other, not always easily comprehended, "in house" priorities. It has also been annoying to base whole procedures on a technical constraint, (such as a separate sign-up to the system to produce a catalogue of a subset of the collection), only to find these redundant when that constraint is lifted, sometimes without much notice that this will occur;

fifth, it has been necessary for QIT to plan to provide its own "copy level" online catalogue to provide comprehensive item level information. The network's products are very limiting in providing good copy level information. It is particularly deficient in its ability to handle holdings statements for serial records;

sixth, the system is complex, and considerable training costs are involved in the recruiting of staff, at whatever level, and regardless of experience of other online systems;

seventh, ABN is a large network, catering for a very wide range of members. Initial concerns QIT had about a general lack of satisfactory control over the general direction of the network have remained, although these are often manifest as a "feeling" rather than a practical problem. Nevertheless, it is fair to say that member libraries have often felt the network as a whole to be unresponsive to certain priorities, unable through sheer operational load to deal with a range of user queries and inputs, and have sometimes felt the ABN explanations of this slowness to be unsatisfactory. As happens with other software primarily developed overseas, Australian users soon face the dilemma of deciding whether to strike out on their own enhancement path and risk not being readily able to pick up further development by the overseas developer, or whether to inhibit "Australianisation" in the interests of staying in line with the source product. This has been true of ABN and its relationship to the WLN (originally Washington Library Network) software.

What this general concern about control ultimately leads to, as it must, is a preoccupation with the systems of advice to and governance of, the network. Because ABN has become so large, and now has five categories of user without having launched major subsystems (such as acquisitions and interlibrary loan), these concerns have continued rather than abated. It is, from the points of view both of the members and (I'm sure) the National Library, a continuing area of difficulty.

7. Conclusion

Before I sum up, and record a verdict, as it were, I'd like to make one other observation.

Dealing with ABN as an organisation from QIT's point of view has been a quite different experience to that of dealing with other vendors of computer systems which QIT has used. Generally speaking, it has been characterised by greater trust, and greater tolerance on both sides. We, as a "client" are prepared to put up with more frustration than we would normally, because we have an underlying trust in the information we are given to explain a particular problem or delay.

Similarly, ABN puts up with client irregularities (and I include QIT here) with forbearance. In other words, we never experience the insecurity of suspecting that ABN will simply give it all away and leave us stranded, and ABN for its part, can reasonably assume good faith on the part of its users. It is a participative public enterprise (although there will always be discussion about how "participative",, not driven by a board of directors whose only concern is a 40% annual profit margin.

One of the effects of this, from my point of view, is that its future as a part of my library's operations is assured. And this means that we have all the more interest in precisely what form the network will assume in future. This form is made more complex, more uncertain by the changing role and potential of the network.

It commenced life unambiguously as a shared cataloguing facility. Planners saw and hoped that it would become the national bibliographic database. But although decisions have been made based on this assumption (the winding up of the previous gener-

ation of national union catalogues), the network is, in 1988, by no means a national union catalogue. Major library collections, noticeably in my own region, are absent, and show no present prospect of joining.

Debate, sometimes uninformed, continues on what policies should be adopted to strengthen the network as the National Bibliographic Database. Similarly, debate continues on the offering of other services (i.e. for other library functions), in terms of feasibility, costs, etc. I am sure that these debates have been just as vigorous within ABN as in the library community as a whole.

As a member, I continue to be interested in the extension of services. But I have to confess also to a feeling of a certain proprietary stake in all this . . . I become irritated when I hear non-members criticising the unwillingness of ABN to offer greater inducements to join. This unease is intensified by the observation that QIT would presumably help to subsidise such inducements, which were not available when we joined.

Looking back over QIT's experience, and reviewing the benefits and drawbacks of that experience as I have done today, I find without hesitation, that the benefits outweigh the drawbacks. I would draw this conclusion based on cost arguments alone. But the further, clinching factor, is that QIT did, with this step, break the bonds of a professional and functional isolation, perhaps more keenly felt in Australia than many other countries, and now reaps the tangible rewards, every day, of "living together" in this national network.

Thank you.

References

[1] Australian Bureau of Statistics, Australia at a Glance 1987, ABS Catalogue Number 1309.0

[2] Figures from

(a) Australian Bibliographic Network, The Network and the Nation: the Work of ABN 1981–84, Canberra, National Library of Australia, 1984.

(b) Australian Bibliographic Network, Network Committee, Agenda Papers, 24–25 February 1988.

032-CAT-2-E

PICA, FROM PROJECT TO INSTITUTION A PARTICIPANT'S VIEW

Jacques J.M. van Gent, *State University Library, Leiden, The Netherlands*

Cees A. Klapwijk, *State University Library, Leiden, The Netherlands*

ABSTRACT

PICA (Project for Integrated Catalogue Automation) has grown into a library-network with facilities that include the initial shared cataloguing, a union catalogue, an interlibrary loan-system and an online retrieval system, in which nearly all important research libraries, the national Royal Library and the larger public libraries participate. Progress in automation (online techniques!) and shrinking budgets ask for a fresh look at the necessity of maximum level cataloguing but require also local reorganization of all work that deals with information stored on general bibliographic, local bibliographic and copy levels. The authors, going one step further, propose the setting up of an interlibrary agency from which all participants of the network may acquire their book material together with bibliographical data of high-quality standard provided for them in the central database.

History

The first impulse for what was later to grow into the cooperative Project for Integrated Catalogue Automation (PICA) was given by the idea that cataloguing efforts should be shared between all participants.

At the end of the sixties, the time in which PICA was initiated, only a few automated library systems were in operation in various parts of the world, and none of these seemed to have the scope and the quality that were envisaged by the group of cataloguing experts called together and led by G. van Wesemael.

From a quality point of view the shared cataloguing concept was based on the words of Henriette Avram that the minimum degree of completeness of bibliographic data would be determined by what was required for uniqueness in the master database to which all bibliographic records should be related. On that principle the Library of Congress and the British National Bibliography had taken the position that the records distributed should be as rich in detail as possible: the whole idea was, that if the manufacturing of a title was a once only effort, it might as well be done as extensively and detailed as possible.

No need to say that the Van Wesemael group was of the same opinion being a staunch supporter of the Paris Principles: the catalogue should be an efficient instrument for ascertaining whether the library contains a particular book specified by its author, its title or a suitable substitute for both and which works by a particular author and which edition of a particular work are in the library.

It was a time in which the sky was the limit. We, that is the group mentioned above, made plans for the integrated total library systems, the integration being vertically worked out within the local library procedures and horizontally where each library-procedure had to be shared with all other libraries participating in the project. At all times, we could not deny our background: the catalogue and cataloguing were at the center of it all. But it was clear to us, that the shared catalogue and the shared cataloguing resulted, or should result in shared acquisition, shared classification, shared circulation control, all come together in that one and only Dutch total integrated library system PICA.

Indeed the Van Wesemael group had arrived at a state where all plans were of a vision splendid, but at the same time materialization was poor, and time was running out. As the late Mr. De Groot, former chief librarian of Leiden State University stated: "After five years of research, different kinds of problems occurred, the most important of which was the impatience to be able to show tangible results. The results of the research were technically very advanced and specifically user-oriented; but short-term results which should also achieve an improvement in the efficiency of library processes were not in evidence".[1]

A change in approach was decided upon, an approach aimed at a realization in stages, to begin with the shared production of bibliographical records and output products for all types of catalogues.

A number of visits to the United States, where Mr. Kilgour with his OCLC was remarkably successful, contributed greatly to that change and much credit must be given to Van Wesemael that he, be it with pain in heart, decided to leave the work he started to others who were better equipped to make the down to earth first practical steps that were necessary at that time to get things going. His satisfaction must be great to see nowadays how many of his ideas have come to reality. PICA has grown into a library-network with facilities that include the initial shared cataloguing, a union catalogue, an interlibrary-loan-system and an on-line retrieval system, nearly all important research libraries in the Netherlands, the National Royal Library and the larger public libraries participating.

There is no need to stress the importance of the PICA-system and the value of its products to the Dutch library community any further. It suffices to state that the whole two-decades history of PICA shows that the idea was not bad at all, to say the least, and that it worked out quite well.

Shared or cooperative cataloguing

Let us now have a look at some of the special aspects of that one part of the PICA-system that is important to us in this section of IFLA: the shared cataloguing and some of its technical and organizational problems in the Dutch library environment.

The main task of the PICA-database is to support the shared cataloguing system which is a realization of the principle that "a publication in the Netherlands will be catalogued only once and then only if its title is not available from a national bibliography, which data are stored in the central PICA-database". This is a quote from a statement of the Dutch Minister of Education in 1977 on the governmental policy on library automation in the Dutch research libraries.[2]

The statement was a clear indication of the fact that the Department of Education had kept faith in the PICA principle of shared cataloguing and was willing to continue its financial commitment when the system at that time had only just emerged from its growth pangs and was on its first feeble way to adolescence. Theoretical concepts had been transformed to a real and tangible fast growing database fed by off-line input of national bibliographic tapes, from which all participating libraries could use titles found in the database and entered new ones if a title was not found. The combination of hard work and ministerial support had paid off.

The technical and organizational problems mentioned above may be familiar to many of us. However to clarify the situation a few words have to be said about the way cataloguing is done in the PICA system.

Information of general interest as well as information of local

library importance is added to all titles, found or entered in the database. The information is stored on different levels:

— general bibliographic level, containing bibliographic title information which can be shared by all libraries.
— local bibliographic level, containing all data that are only of interest to one specific library, and which are only accessible to the library concerned. These data are valid for all copies of a publication. An example of data on this level is a systematic classification according to a system designed by a library.
— copy level, which is pertinent to the specific copies and can consist of information such as booknumber, annotation, selection-key for off-line output, etc.

For quality control as a part of the general bibliographic level a shared authority-thesaurus is operational and for some purposes on local level, thesauri for corporate authors and subject-headings are being created.

The bibliographic title-information to be found on the first level had to be based on whatever was required by the Dutch Cataloguing rules, which are in many ways a faithful copy of the ISBD's.

PICA had always been of the opinion, and somehow still is, that it would strengthen its position within the Dutch library community by sticking to the rules even though some experts at the PICA-bureau considered them rather cumbersome and in some instances unfit for a proper functioning of the automated system. Painstaking, time consuming work required by the rules and regulations of the ISBD's provided titles as rich in bibliographic detail as possible that would be of benefit to all participants. Why antagonize cataloguers, if, within the system, it was going to be a once only effort?

Minimum level cataloguing

Yet it became necessary for representatives of PICA and the Dutch Federation of Library Organizations, the latter being responsible for the rules, to meet each other in June 1986 in the Royal Library. Minutes of the meeting speak of progress in automation and shrinking budgets that asked for a fresh look at the necessity of maximum level cataloguing. A committee set to work quickly decided that minimum level cataloguing should not be taken to imply minimum quality, this being the same view as taken by the British Library in its cataloguing action plan, but also that the type of catalogue it was aiming for should not be considered to act as a bibliographic tool. In other words, the Paris Principles had to give way with regards to the principle that a catalogue should tell which works by a particular author and which edition of a particular work are in the library. The problems are neither particularly Dutch nor of recent times only. With regards to maximum cataloguing it suffices to quote Edmund Lester Pearson, a man who seemed to have a natural antipathy to cataloguers of all varieties who said in 1914 "the man who wants to know about important books, ancient or recent, will insist on seeing the books themselves. No catalogue entries will satisfy him. The man who does not care to know about important books does not need an elaborate catalogue — and so, who in the devil does need an elaborate catalogue?"[3]

Undoubtedly production levels have gone up since 1947 when the library of Congress estimates spoke of an average time required to perform descriptive cataloguing being two hours twenty minutes per title, or three books per cataloguer in a seven hours day. However there must be plenty of reasons to still harp on the subject considering the fact that D.J. Urquhart's statement that "delays in making items available, even if they result in better records, may in fact produce a decrease in effectiveness and an increase in costs — a particular example of the general principle that the best is the enemy of the good", finds a prominent place in the consultative paper "Currency with coverage" of the British Library.[4]

With regards to the Paris Principles it may be clear that dropping the necessity to look for the main entry heading, to mention one item only, the discussion of which about its concept has occupied more of our time to little purpose than everything else in librarianship, as A.J. Wells ventured to say more than 10 years ago,[5] should save a lot of work. Time will prove if the work carried out by the committee, that resulted in a new set of minimum level cataloguing rules will lead to the better results that those who were responsible for this initiative, aimed at.

Local organization

Cataloguing, especially cooperative or shared cataloguing however is not only a matter of (cataloguing) technique but also a matter of organization of library work. PICA claims, with figures to prove it, that the libraries that participate in PICA will come into immediate reach of standard bibliographic records of nearly 80% of all required titles. Why then, one could ask, is it necessary to go look for new minimum level cataloguing rules and run the risk of endangering the quality of the information in the database, if only 20% of all titles need to be catalogued as new input?

More attention should be paid to the fact that each good title, once added to the database does not need to be handled anymore. Not cataloguing rules slow down production but such organization of work that allows each title to be handled over and over again. Paradoxically only complete, high-standard titles, added to the database on a maximum level cataloguing basis may improve production, because their guaranteed quality will be accepted by each library without any further time consuming control functions.

It may seem reasonable then that books of which bibliographic descriptions have been found in the database will only be handled by those library departments that are functionally equipped to deal with the storage of information that is necessary for the local and copy level and not by the cataloguing-department at all, the latter being responsible for the general bibliographic information only, which means the remaining 20% of the titles not found in the database.

The original concept of vertical and horizontal integration will then be the basis from which internal and interlibrary organization could be realized. Within the libraries cataloguers should be encouraged not to keep the central database to themselves. Librarians should enforce the rule that each input that contributes to the completeness of bibliographic data on the local and copy level will be handled by the appropriate departments, like research workers signaling their desiderata straight into the database, and the acquisition department adding its acquisition data to the title involved. At the same time, to ensure quality and a minimum degree of completeness of bibliographic data in the master database, librarians should agree upon which titles are going to be handled by which libraries in such a way that the resulting qualified input will always be accepted, without time consuming control by others.

Regional or national organization

In spite of the infrastructure created by PICA there still seems to be a considerable overlap of acquisition and cataloguing work, not only, as we have seen within each library separately but between all participating libraries as well. With regard to the necessary internal reorganization libraries have to work things out for themselves, like indicated above, if they want to make full use of the PICA-system. With regards to reorganization on a national scale a lot depends upon the secured quality of bibliographic data added to and procured from the central database. If libraries are unable to set up a scheme of commitments for participating libraries to add high standard entries belonging to their own appointed field or to change titles accordingly the following suggestion could be considered. Some kind of an agency, call it a "book/title shop" should be set up by the Dutch universities from which all university libraries are obliged to acquire their material. The very fact that PICA covers just about all research/university libraries needs, makes this suggestion all the more obvious. A tremendous logistical advantage may be gained over the present agreements to buy and catalogue cooperatively on an voluntary basis. If libraries ensure themselves of acquiring book material together with bibliographical data of high-quality standard, doubtless the sum total of publisher's price, plus acquisitions cost plus cataloguing cost will be considerably less than the cost of acquiring and cataloguing books if acquired and catalogued by each library separately. Non participating libraries may buy this service also and get themselves provided with any type of catalogue output they prefer. The attraction of it all may even be enhanced by the possibilities of shared classification input either by the participants or, better still, by the agency itself. Here too, coordinated collectionbuilding may be considered to be practical at last. The suggestion seems to be worthwhile enough to study it seriously.

BIBLIOGRAPHY

1. J.R. de Groot. "PICA and the humanities". In: Large Libraries and new technological developments. Proceedings of a symposium held on the occasion of the inauguration of the new building of the Royal Library, The Hague, 29 September–1 October 1982. Ed. by C. Reedijk e.a. Munchen, Saur, 1984. p. 35–42.
2. Anton Bossers and Martin van Muyen. "Library automation in the Netherlands and PICA." In: The Electronic Library 2, no 2 (1984), pp. 87–99.
3. K.D. Metcalf. The attitude of the library administrator toward cataloging. In Catalogers' and classifiers' Yearbook 10 (1941) pp. 9–18.
4. Currency with coverage. The future development of the British National Bibliographic Service. Consultative paper. In: The British Library Bibliographic Services. London, July 1987.

[5] A.J. Wells. The British national bibliography. In: The Brasenose conference on the automation of libraries; proceedings of the Anglo-American conference on the mechanization of library services; University of Oxford 1966. London, Mansell, 1967.

021-CLASS-1-E

CLASSIFICATION ONLINE: PRESENT AND FUTURE

Nancy J. Williamson, *Faculty of Library and Information Science University of Toronto Toronto, Canada*

ABSTRACT

This third report from a project sponsored by the IFLA Section on Classification and Indexing focuses upon the Library of Congress Classification. With the Dewey Decimal Classification project as a partial model, Dr. Nancy Williamson, Project Director, received an invitation from the Library of Congress to study the online possibilities of LCC. A tentative MARC record format for LCC was tested against a 4,000 page random sample from the LCC schedules. Preliminary findings indicate that a MARC record format is possible, especially after rationalization of LCC by consistent editing. Two major decision areas will be the handling of hierarchies of terms and the treatment of tables. Small scale tests suggest the possibility of extending this format to other major classifications schemes with the eventual goal of an international MARC format for classification.

Introduction

In 1984, the IFLA Standing Committee on Classification and Indexing, endorsed a project for the study of the role of classification and classificatory structure in all kinds of online retrieval systems. The overall objective of the study has been to address the question 'What is the place of classification in online systems?', in order to identify and define the research needed to arrive at possible answers to this question. Also it has been the intent to identify some projects which could suitably be fostered and undertaken by the Section.

The investigation has focussed on three crucial areas — online catalogues, bibliographic databases and full-text information systems. The early stages of this study were described in two papers presented at the IFLA Conference in Chicago[1] and Tokyo.[2] The discussion which follows here, briefly describes investigation and research undertaken since Tokyo and summarizes the findings, conclusions and recommendations of one aspect of the study — classification in online catalogues. Because of space limitations it was not possible to cover all three aspects of the study in one paper. However, more detailed comprehensive reports on all three aspects of the study are being submitted to the Standing Committee at the Sydney Conference.

Classification and the Online Catalogue

Because catalogue records usually contain classification numbers assigned to represent the subject content of documents, online catalogues are obvious vehicles for research on book classification and its role in online bibliographic retrieval systems. Historically some groundwork has already been laid and there are online catalogues readily available on which to carry out research and investigation.

Historical Context

Interest in the mechanization of book classification is not new. It can be traced back at least 40 years to numerous projects which are well documented in the literature.[3,4,5] At a very early stage the Universal Decimal Classification (UDC) was recognized as being amenable to mechanized sorting. Unfortunately, the earliest research, carried out when computer technology was still primitive, was not encouraging. There were those who declared that UDC was simply "unmachinable".[6] Fortunately, this did not deter more eager researchers and many projects were launched between 1948 and 1980. Research was, to some extent, isolated and extremely varied in nature. Projects ranged from the maintenance and display of schedules, through the indexing of bibliographic files and vocabulary control and thesaurus construction, to the display of bibliographic records and the use of classification in interactive retrieval. Most of the projects predated online catalogues. Some were experimental, while others were applications to operational systems, some of which are still in existence. Most closely related to recent research was the development and experimentation with the online interactive retrieval system called AUDACIOUS[7] by Robert Freeman and Pauline Atherton Cochrane. This study used UDC schedules for browsing and for organizing bibliographic files in the nuclear sciences.

While not all of the research used traditional library classification, most projects centered on UDC as the first classification system available in machine-readable form. Logically then, we might have expected UDC to leap into the latter part of the 20th century, maintaining its lead in mechanization. Strangely, this has not happened. Financial and administrative problems have intervened. Machine-readable files of UDC have been used by individual libraries to create classified catalogues but the system available is in no way as sophisticated and useful as it could be with the present state of technology. However, in 1987 there was a positive move when a UDC Computerization Working Party was appointed "to set up a system to facilitate the revision, publication and computerized application of UDC."[8] Moreover, even though UDC has not moved with the times, the early research has been invaluable as a lead-in to classification research on online catalogues in the 1980's.

The Dewey Decimal Classification project,[9] completed in 1985, which drew on the AUDACIOUS project for inspiration, is one of the most significant research projects in online retrieval in recent years. Sponsored by the OCLC Online Computer Library Center, the Forest Press and the Council on Library Resources, the research focussed on classification as a user's tool in the online catalogue. Its objectives were the study of strategies for searching and display of classification schedules online and the demonstration, testing and evaluation of a classification scheme as a searcher's tool. The findings were both interesting and valuable, and a warning to future systems designers "not (to) repeat the original design and implementation but (to) follow ... recommendations for improving searches[10] is extremely important. This study provided a great deal of specific data on subject retrieval from online catalogues, but it also emphasized some general principles of information retrieval which are essential to systems design. It demonstrated the fact that recall can be improved by providing more than one mode of subject access. Also the failure analysis reaffirmed the importance of the matching process in retrieval. That which can be retrieved is that which has been indexed (or in this case classified). Cutter said "enter a work under its subject heading, not under the heading of a class which includes that subject."[11] This principle of specificity can just as well be applied to the process of classification. In the bibliographic online world, the application of this principle is still necessary if there is to be precision in the retrieval from online catalogues.

The Library of Congress Classification Project

a) Introduction

In 1987 the Library of Congress (LC) took its first steps toward the automation of classification at LC. It is a project which presents enormous challenges, but senior staff at LC recognize that the time has arrived for the development of an online system for the Library of Congress Classification (LCC). Moreover, encouragement has come from several quarters. The DDC project has demonstrated the value of classification in online catalogues. Cochrane and Markey[12] predicted "that someday there will be a machine-readable version of the Library of Congress classification" and analysis of problems and prospects for LCC by Chan[13] and Williamson[14] suggested that such a goal might be attainable.

Accordingly the wheels have been set in motion. By November 1987 discussions had resulted in a set of long range objectives and a tentative plan had been developed for an analysis of the printed schedules and the preparation of a profile of LCC to lay the groundwork for the project. Subsequently, at the American Library Association Midwinter Meeting in January 1988 it was announced that LC would be proceeding with the mechanization of its classification, that LC staff had been assigned to work on the project and that Nancy Williamson would carry out the basic research necessary to provide the data required to proceed with the conversion of LCC and its indexes for use in an online interactive system.

It is LC's intention to design an integrated system which, eventually, will perform a number of functions and which will meet four clearly defined goals.

1. An online interactive support system to enable LC to edit, update and maintain its classification schedules online;
2. A system which will permit classifiers at LC and in other libraries to classify and shelf-list online;
3. The production of the LCC schedules in a variety of physical forms — paper volumes, computer tape, CD-ROM, etc., and

4. A catalogue users search tool which would enhance subject retrieval in online catalogues.

Such a system assumes links among classification schedules, bibliographic records, a shelf list and an authority control system.

A system of this magnitude and complexity must necessarily evolve one function at a time, but if success is to be achieved, careful research and planning at the outset are essential. With a 'collection of systems' such as LCC it is important to be doubly careful. Inconsistent and unusual data must be identified, categorized and reconciled before the system can be designed. Hence the need for a detailed analysis and profile of the schedules.

b) A "MARC" format for classification

The first step in the research was the development of a MARC record format for classification at LC. Long experience with machine-readable data by LC staff led to the conclusion that the most expedient approach to the research would be to begin with a record format, however tentative it might be. Data elements from the schedules could then be related to the format and the format revised as the research progressed. This has proved to be an extremely wise decision. The format, which can best be described as 'a working format for purposes of research', has served a number of important functions in the process of research. It has provided a focus for the analysis and has been essential in the establishment of definitions and in identifying and categorizing unusual features of LCC.

The working format parallels the US MARC format for bibliographic records and in a single record provides for all of the data elements so far identified from the schedules, indexes and tables. The intellectual content of the working format includes the following data elements:

Classification notations (from schedules and tables)
Captions (from schedules and tables)
Table numbers
See references
Notes (identified by function, e.g. scope notes class elsewhere notes, subarrangement instructions, etc.)
Preceding hierarchy
Index terms

A particular record is unlikely to have all of these elements, but every record must have a minimum of two elements — a notation and a caption. Two of the most crucial issues in connection with the format were the handling of hierarchical relationships and index terms. Since LCC has a non-hierarchical notation, two other approaches were considered. Either the records could be linked or the hierarchies could be stored with each record. Initially, the latter course has been chosen for two reasons. The linking of records would greatly increase the complexity of the system and could also slow the process of display significantly. In the case of the index terms, it would be possible either to create separate indexes or to store the index terms with the captions to which they refer. It was decided to store the index terms with the captions. Indexes could be created from the master record. Also, from the point of view of editing and maintenance, there are certain advantages to being able to display in one record all the data elements related to a particular caption. Of course, a window display might be another solution.

c) Data collection and analysis

In preparation for the data collection decisions were needed on what data to collect and how they would be used. Using the working format, randomly selected pages from such diverse schedules as N (Fine Arts), PQ, Part 2 (Italian, Spanish and Portuguese Literature), B-BJ (Philosophy, Psychology), KK-KKC (Law of Germany) and HM-HX (Social Sciences: Sociology) were examined and several hundred records were coded. While time-consuming, this primary study served to formulate the questions to be addressed in the data collection and to reveal potential problems in developing an official US MARC format. Clearly both qualitative and quantitative data needed to be collected, and any unusual data needed to be categorized as to its nature and function.

When the proposal for data collection was complete, the requirements were discussed with LC's statistician for advice on the sampling technique to be used. The decision was to sample all 41 schedules, in each case using 100 randomly selected pages for a total of approximately 4,000 pages. Every piece of data on a page was to be scrutined, characterized and counted and data recorded on a two-page sampling form prepared for the purpose. Notes would be collected on unusual features of the data. Also to be collected was any information about data which would affect the editing process when the schedules are eventually converted to machine-readable form. It was important to use the most complete and most up-to-date schedules available for the analysis. Therefore, very early in the discussion, it was decided that the Gale Research Company's cumulated edition of the schedules would be used to develop the profile. Gale Research generously agreed to supply this edition of LCC for purpose of the research.

Following these discussions, the sampling and data gathering techniques were pre-tested across a wide range of the schedules. This served to clarify any problems and to identify minor adjustments which were needed in the sampling methods. The last stage of the project, still in progress, is the collection of the data from 4,000 pages of LCC and their analysis. Computer programs have been designed to aid in the analysis and the data collected is being converted into machine-readable form as it is being collected. Status reports are available as the project proceeds.

d) Findings and preliminary conclusions

At the time of writing, it is premature to state precisely what the final results will be. Nevertheless, certain trends are clearly evident from the preliminary study, the pre-test and the final data analysis, insofar as the latter has been completed.

Prior to beginning the project, there was reason to seriously question whether the long term goals could be accomplished. In the early stages, random examination of pages from the LCC schedules revealed numerous problems to be solved and challenges to be met. However, gradually the volume of new "discoveries" has decreased and the signs have become both encouraging and exciting.

There is now a substantial degree of certainty that a satisfactory MARC record format can be achieved for the LC classification. While the officially approved format may differ substantially for the 'working format' used in the research, if a MARC format is possible, a major step in the direction of the long-term goals has been achieved. Careful scrutiny of each piece of data sampled will remain important throughout the project. As data collection proceeds, there are fewer data elements being located which are actually new data types. Nevertheless, as a book cannot always be judged by its cover, so also new LCC data may not always be what it seems. Even the slightest variation in wording must instinctively be questioned. What kind of data element is this? What function does it perform here? Can it be categorized with other data types already identified and defined? Or, is it something entirely new? Sometimes variations are very subtle. More frequently than not, it is the wording of notes and instructions which varies and when the intent is analyzed these notes fit into the categories of data elements already identified. For example, notes which read "In case of doubt see JS" and "For history of the science see HVG021" both function as class elsewhere notes, while f.i. (for instance) is really an archaic form of e.g. (for example). Variations and inconsistencies such as these occur most frequently across schedules, rather than within a single schedule. The publication date of a schedule is also a significant factor. A schedule published in 1980 will be at greater variance with a schedule issued in 1960, than in the case with schedules published in the same year. There are obvious typographical errors and a fair number of curiosities (e.g. Marx, Karl see Karl Marx). All of these things are candidates for rationalizing and editing data when the schedules are converted to machine-readable form, but they are not obstacles to the development of a MARC format or to the mechanization of the system. In the overview there is surprising uniformity in the intent of notes and instructions across schedules and it is not unreasonable to expect that the completed analysis will show that the number of these exceptions to the rule may be small in comparison with the total number of records in the schedules.

The sampling of the LCC schedules will yield data which will permit the research team to quantify the approximate number of records to be contained in the database and the average number of characters in each record as well as to identify types and quantities of notes, references and instructions. Since every piece of data on a page sampled is being scrutinized, a by-product of the quantitative approach is the identification of precise errors, omissions and inconsistencies as well as unusual and unique data. As might be expected, three important aspects of LCC requiring special attention — the hierarchies, the tables and the divide-like instructions.

The handling of the hierarchies of terms is of major concern, for at least two reasons. There is the absence of a hierarchical notation in LCC, and many captions, such as "General Works", "Periodicals" and, above all, "Other", are completely lacking in expressiveness. Furthermore terms such as "Other" are meaningless in a hierarchy unless the display allows the user to determine "other than what?" Some such terms will have to be edited into something

more meaningful if they are to serve any function at all. There are at least two possible methods for providing the hierarchical relations in this situation — physical linking of records, or the storing of the hierarchy for each topic with its caption.

There are advantages and disadvantages to both approaches. However, the basic question which must be answered is "Which is the best way to accomplish this? Factors which will play a part in answering this question are complexity, the effect on the size of the records, the size of the database and the requirements for maintenance, manipulation and display of the schedules. The sampling process allows for identification of up to 20 levels in a hierarchy. As yet no 20-level hierarchy has been located. Up to this point in the data collection the 'winner' is KK-KKC (Law of Germany) which displays up to 13 levels of hierarchy. Fortunately an extremely high number of hierarchical levels is proving to be the exception rather than the norm. Excessively long hierarchies are potential problems in screen display.

The life blood of the LCC is its tables. There are large tables and small tables, simple tables are complex tables, tables within tables, and tables which are same as other tables. Particularly complex are the large tables, embedded in the main schedules, which often require classifiers to create their class numbers by moving from schedule to table and to a second table, and sometimes even on to a third table. The use of a MARC format to encode LCC tables does not present serious problems, because many of the data elements in tables are the same data elements as found in the main schedules. However, in moving into an an online interactive system, one of the goals should be to simplify the manipulation of the tables if possible. There are a number of questions being asked. Could we eliminate some of the tables? Possibly. Certainly, there are a number of tables, particularly small tables, which duplicate each other, but it would also be helpful if some of the time-consuming manipulation of the more complex tables could be either removed or reduced. Would it be useful to extract all of the tables and put them together in the database and in a separate printed volume? A useful display in the database may not satisfy the needs of users of the printed volumes. It is very useful to have small tables on the same page of the schedules as the class numbers with which you are now working. Since both physical formats will be created from the same database, this will need to be taken into account. Thus many questions in connection with tables need further study, but it is apparent that further research on the tables must be cognizant of two things — the way in which those tables will be used in different physical formats, and careful investigation of the capabilities of the technology. It is not inconceivable that algorithms could be developed so that, on a given set of commands, class numbers could be created through machine manipulation of tables, rather than human calculation.

Divide-like instructions, appear frequently in some schedules (e.g. N Fine Arts) and not at all in others. The divide-like instruction is a space saving device for printed schedules. In an online system, this same kind of space saving is not required. The data analysis is designed to provide answers to such questions as "Approximately how many divide-like instructions are there? How many class numbers are encompassed by these instructions? Would it be useful to print these numbers out in the schedules, rather than using this device? Answers to these questions depend on the effect it would have on the potential size of the database and, more importantly, on the impact on a printed version of the schedules. Depending on the answers to these questions the divide-like instructions could disappear altogether from the schedules.

f) Implications of the research

The analysis of the schedules is already having an impact at the Library of Congress. Discussions have taken on an atmosphere of optimism and excitement and LC staff and the research team are looking at the schedules in a new light. There is an increasing awareness of inconsistencies and possibilities for improvement. The list of data elements to be edited is increasing almost daily and new policies are emerging. The cataloguing world in general will probably first see the impact of the data analysis in new editions of the printed schedules, published within the next year. For example there may be changes in the way the A-Z instructions are handled. Also, captions such as "other" may be replaced by more explicit terminology. Upon completion of the data analysis, a period of decision-making and testing is required to finally determine the MARC record format. Only then can the conversion of the LCC schedules begin. While the process appears to be long, there is a definite commitment and work has begun.

Further Research

If a satisfactory MARC record format can be developed for the Library of Congress Classification, it is reasonable to assume that the same can be done for other classification systems. As an extension of the LC project, the "working format" has been applied, with success, to a sample of entries from the 19th edition of the Dewey Decimal Classification, the UDC International Medium Edition (English Text) and the Bliss Bibliographic Classification. In each case a more in-depth study is needed, but the time is right to begin to consider the possibility of an international MARC format which could accommodate the world's major schemes. We cannot know with certainty whether such a format could be developed until more research is done. There is an opportunity here for the Standing Committee on Classification and Indexing to foster research which is of great importance to the world of classification.

References

1. Nancy J. Williamson, "Classification in Online Systems: Research and the North American Perspective", International Cataloguing (July-September 1985): 29-31.
2. Nancy J. Williamson, "The Library of Congress Classification: Problems and Prospects in Online Retrieval", International Cataloguing (October-December 1986): 45-48.
3. Malcolm Rigby, Computers and the UDC: A Decade of Progress, 1963-73. (The Hague: International Federation For Documentation, 1974).
4. Malcolm Rigby, Automation and the UDC, 1948-1980 2nd ed. (The Hague: International Federation for Documentation, 1981) Pg. 17-30.
5. Pauline A. Cochrane and Karen Markey, "Preparing For the Use of Classification in Online Cataloging Systems and in Online Catalogs", Information Technology and Libraries 4 (June 1985): 91-111.
6. Barbara Kyle, "Merits and Demerits of Various Classification Schemes for the Social Sciences", Unesco Bulletin 14 (March 1960): 54-60.
7. Robert R. Freeman and Pauline Atherton, "AUDACIOUS: An Experiment with an Online Interactive Reference Retrieval System Using the Universal Decimal Classification as the Indexing Language for the Field of Nuclear Science." Information Transfer: Proceedings of the American Society for Information Science. Annual Meeting, Columbus, Ohio, Oct. 20-24, 1968. (New York: Greenwood Press, 1968): V.5, pp. 193-199.
8. "UDC Computerization Working Party", Information Hotline 6 (March 1987): 12.
9. Karen Markey and Anh N. Demeyer, Dewey Decimal Classification Online Project. Dublin, Ohio: OCLC Online Computer Library Center, 1986.
10. Karen Markey, "Searching and Browsing the Dewey Decimal Classification in an Online Catalog", Cataloging & Classification Quarterly 7 (Spring 1987): p.68.
11. Charles A. Cutter, Rules for a Dictionary Catalog, 4th ed. rewritten. (Washington: US Bureau of Education, 1904): p.66.
12. Cochrane and Markey, p.91.
13. Lois Mai Chan, "Library of Congress classification as an Online Retrieval Tool: Potentials and Limitations", Information Technology and Libraries 5 (September 1986): 181-192.
14. Williamson, "The Library of Congress Classification: Problems and Prospects in Online Retrieval".

022-CLASS-2-E

AUSTRALIAN DIRECTIONS IN SUBJECT ACCESS

Elaine Hall, *State Library of New South Wales, Sydney, New South Wales, Australia*

ABSTRACT

The majority of Australian libraries base their subject cataloguing on Library of Congress Subject Headings (LCSH). To assist with the application of LCSH, and to provide guidance on building subject access on the Australian Bibliographic Network database, the Australian subject cataloguing manual was published in 1985. Aspects of the suitability of LCSH for Australian libraries are covered. However, in an environment where there is much dissatisfaction with the quality of access being provided, a change of direction is proposed, involving a comprehensive evaluation of the appropriateness, effectiveness and efficiency of the subject access process.

I devised the title to this paper before I wrote it, and in response to an invitation to speak on various aspects of subject access in the

country hosting the IFLA Conference at which the paper is to be presented. In the course of thinking further on the topic and actually writing the paper, I became increasingly concerned about what may be perceived as the presumptuousness of the title. So I hasten to make it clear that the paper is essentially a personal view of some recent directions in subject access in Australia; also it is not an historical overview.

Firstly I propose to highlight the major directions that have been taken in subject access in Australia in the last few years. This will lead on to some changes in direction, and the directions that I think should, and indeed must be considered for the future.

In Australian libraries, subject access is to be equated with subject cataloguing using Library of Congress Subject Headings (LCSH). Much as any sweeping statement is not entirely true, the comprehensive accuracy of this statement can readily be challenged. Nonetheless, the bulk of cataloguers around Australia use LCSH, or base their subject cataloguing on LCSH. Now whether the Library of Congress would always recognise the subject headings produced in all these libraries as LCSH is a question that I will pose, without pursuing the answer.

As an aside at this point, it is interesting to note that Australia did not, at least in this regard, follow the example of her former colonial power and pursue the classified subject approach. While there were some excursions down that route, and while a sophisticated classified approach may be the practice for some specialised collections, the predominant general library tradition has been that of the alphabetical subject catalogue.

But what does it mean to say that you "use LCSH?" My tendency to ask such questions reached I think its zenith in 1983, in a paper to the Australian Bibliographic Network (ABN) Standards Committee. In that paper I bemoaned that the problem was that there was no AACR2 of subject cataloguing. The outcome was that I undertook, with very considerable assistance from the National Library of Australia, and in particular from the staff of the ABN Office, the compilation of the ABN subject cataloguing manual.[1] The manual, published in 1985 was directed to providing guidance in building subject access on the ABN database.

At the same time, at first unknown, the Library of Congress was preparing the preliminary edition of their subject cataloguing manual. In a nutshell, the ABN manual presents the same material as the Library of Congress manual, excluding the details of Library of Congress procedures, and including local practices as approved by the ABN Standards Committee for the ABN database.

Certainly, not comprehensively, but very extensively, the manual tells you what "using LCSH" is about — the practice of applying LCSH. But of course not the theory, nor the philosophy — there is essentially no theory or philosophy, just practice. And this is where a blanket of dissatisfaction can begin to form over the satisfaction of knowing how to construct legitimate headings and how to employ standard practices for particular types of material.

It is this dissatisfaction that I will explore further later. But firstly I want to follow some parallel issues.

On the Australian bibliographic horizon, and indeed on the broader Australian library horizon to a significant extent, the brightest star over the last seven years has been the Australian Bibliographic Network — an Australia-wide network of about 140 libraries building a database, currently made up of approximately five and three-quarter million bibliographic records, with almost seven million Australian holdings recorded against them; a system that uses the WLN software. It is always dangerous to claim uniqueness on a world-wide scale, but I think it is true to say that ABN has been the first database to be built using MARC records not only from local cataloguing but also from the national agencies of five countries, with data now 20 years old, and with authority control attempted over the entire database. Now seven years old, the state of subject access on ABN — its consistency, its reference structure, its up-to-dateness — would be an ideal subject for a case study of the resources necessary to maintain a quality database.

One issue in relation to subject access on ABN is of outstanding significance; that is the issue of Australian terminology — both for things unique to Australia, and for Australian usage, both of things Australian and of things from the Australian perspective. In Australia this issue, from its variety of perspectives, has had the highest profile in the criticism of the adoption of Library of Congress Subject Headings in Australian libraries — it has ranked higher than the issues of prejudice in headings, and out-of-dateness. And the resolution of the problems associated with this issue has certainly not satisfied all the people all the time — the cataloguers, to say nothing of the catalogue users.

The argument that Library of Congress headings for Australian subjects that are of great importance will be changed on ABN lead to the decision to change Aborigines, Australian to Australian Aborigines. But "great" is subject to widely varying interpretation. Arguments abound about catalogue users having to be conscious of national differences in usage, in the same way as readers of books from overseas have to be able to "translate" concepts into their terminology. But should Australians have to use Longshoremen instead of Waterside workers or Wharf labourers. And, in an automated environment of increasing sophistication, will the reference structure not provide us with the facility of transparency to circumvent such dilemmas which currently present themselves as issues involving not inconsiderable manual effort to reconfigure. And as an example of a heading which for an Australian user would just simply not be a sought heading, there are the headings beginning East Indian . . . — for the Indians of the Indian subcontinent — which of course happens to be most directly west of Australia. What to do? What has been done? If you were to consult the ABN database you would realise that a number of Australian cataloguers have never been conscious that there is any problem with the use of the word "Indian;" but there certainly are a variety of places to look for a cookbook for example that covers the national dishes of vindaloo, pilau and chappati.

Concerns about headings for things uniquely Australian that are not yet represented in LCSH have to a very large extent been resolved, with the background information on how to establish valid headings, for example for animal and plant names, and with the establishment of a most constructive working relationship between the National Library of Australia and the Library of Congress Subject Cataloguing Division.

Has the issue of the Australianness of LCSH, and hence its fundamental appropriateness as the basis for subject access in Australian libraries, consumed a disproportionate amount of time, energy and resources? While not belittling the place of terminology as an important issue when assessing effective subject access, is it as significant as the resources devoted to it? Is it an issue that, being tangible, could be focussed on? But is it also a classic example of what cataloguers have a bad reputation for — focusing on detail, when the wider issues are more important. Because at the end of the day, there is not a total solution, and the scope of the problem is not as enormous as one might have been led to believe; and most importantly there are much more fundamental issues for the quality of subject access.

What then is the overall state of subject access in Australian libraries today. While it could hardly be contended that it is in blooming good health, there is much activity of interest; for example, there is the developmental work of the Australian Schools Catalogue Information Service, with its ASCIS Subject Headings,[2] the first national list specifically developed for the use of Australian school children. There is the experimental work of the subject augmentation project at the Australian Defence Force Academy Library,[3] a project based on the Subject Access Project devised at Syracuse University. And there is the continuing solid work on the ABN database, where the aim is a consistent, up-to-date, fully referenced authority file. Some may argue that these examples do not adequately represent the range of activity in the field of subject access in Australia at present, but they are outstanding examples. It is symptomatic of the field of activity and interest in the field however that all three are examples of practical applications, where practitioners are trying to push ahead with improving subject access. In other words, the field of subject access has not excited a great deal of interest among academic researchers, at least to the extent of published research. Also the Australian Committee on Cataloguing, a committee formed in 1981 with representatives from the National Library of Australia, and members nominated by the Library Association of Australia and the Australian Advisory Council on Bibliographical Services, with a brief to oversee developments both national and international, in the broad area of cataloguing, has had its attention almost exclusively dominated by revisions to the Anglo-American Cataloguing Rules; and it has only recently begun to get into one area of subject access, namely revisions to the Dewey Decimal Classification schedules.

So that, in the best Australian tradition, the developments in Australian subject access have been very practical exercises. To improve the health of Australian subject access, which I have classed as in a poor state, what direction should be taken. Should it be towards developing a theory and principles for subject access. Should one explore whether one can throw off this blanket of dissatisfaction I alluded to earlier, that may descend when one learns how to do "it," but one still doesn't know why, or to what end. It is interesting to note that the push for developing a theory and a code for LCSH still has its supporters in the United States. I however propose an alternative direction for subject access in Australia. And the first step is to turn aside from any dissatisfaction with not

knowing the whys and wherefores of what we are doing using LCSH.

Next, I think it is vital that we take securely on board the correct enunciation of the fundamental objective of cataloguing departments and cataloguers. That objective must not be defined as the production of a catalogue. Such a definition places unjustifiable emphasis on the process as if it is an end in itself, rather than emphasising the purpose of the process. The objective of cataloguing is to provide access to the wealth of information contained in a library's collection. The aim of a cataloguing section should be to take prime — not sole — responsibility for clients successfully accessing information in a library's collection. This is not a game of semantics; it is a very serious stand that may determine the continuing viability of subject access as an effective service provision in Australian libraries.

Concentration on the catalogue production stance has the tendency to advance the paramountcy of consistency — that all works in the collection must be catalogued at the same level of detail, consistently — an impossible goal at any stage, let alone over time, and the tendency to contend that the catalogue through intelligent exploitation of the reference structure can reveal what the library holds on any subject. A cataloguer will accept that you do not give subject headings to individual works of adult fiction and ignore requests to assist catalogue users by supplying such access. A cataloguer might believe reports that the subject catalogue is not as heavily used as the author-title catalogue without asking why, or troubling to learn that user studies consistently report that the majority of user requests are for material about a subject. This may appear hypercritical of cataloguers; in that case I am being hypercritical of the role I performed in building a catalogue for a number of years. It is meant to draw attention to a few of the major stumbling blocks in providing effective subject access and to provide me with a platform on which to build a scenario for improved subject access. It is not to deny that a soundly built catalogue is a firm foundation onto which improved subject access can be grafted.

The major characteristics of the direction that subject access has been taking over the last several years have been those of sharing and standardisation. Two factors are increasingly forcing a reassessment of direction. Firstly more rigorous management practices are being adopted in a climate witnessing an upsurge in the scrutiny of public institutions and libraries. And secondly, different and incompatible computer systems are one of the major reasons behind the implementation of online public access catalogues along very institution-specific lines.

In an environment where all services are being evaluated, how well would subject access stand up to an in-depth evaluation of its appropriateness, effectiveness and efficiency? How well could one justify the outcomes of the activity — that is, the effective subject access — in terms of the input of staff and other resources? What has the institution determined to be appropriate subject access for its collection, given its overall goals, the nature of its collection and its target client groups?

Evaluation of subject access with a view to improving that access, and at the same time improving the efficiency both in creating the access and for the clients accessing the information, must take into consideration the whole spectrum of the factors that go into providing the current service and the range of possible and even improbable options the institution has at its disposal for improvement. This is not the place for counting on simplistic solutions, such as the employment of additional cataloguers. Nor will the solution simply be found by pouring more money into automation.

The solution to the problem of how to provide appropriate, effective and efficient subject access will come out of considering factors such as the role of the computer — not only its capacity for powerful searching in terms of high volume high speed processing and sophisticated techniques of combining and refining, but also for example its ability to guide and teach can be explored, its capacity for transaction logging, along with its disadvantages such as its unforgivingness and still daunting power must be explored.

There is the role of library staff in assisting clients to be considered; reference staff have of course always served as guides to clients in the use of the catalogue. But would the most efficient way to improve subject access be to provide additional staff to assist and train clients; staff who not only know the collection but also understand thoroughly the way in which the major in-house means of access, the catalogue, has been built, plus know alternative printed and on-line means of access to information. Such staff could feed information from their experience into building the improved access tool, test that, provide feedback. Such staff should also play a vital role in breaking down the barriers between reader services and cataloguing sections; barriers which can be so destructive of moulding the access tool to the access requirements of clients.

There is of course also the role of the catalogue record to be considered. Among a host of factors for analysis is the role of alternative or supplementary modes of access, alternative to the standard subject headings. Or should the standard shared data be modified — simplified for in-house use? What is the real place of a reference structure? And should traditional controlled term access be augmented with uncontrolled terms? Is there a new role for classification in subject access?

Is such a process of evaluation realistic, and will an answer be found? As already indicated evaluation is a necessary element of management, and perhaps even survival. And progress has to be made towards improving subject access. There is no room for exasperating dead ends. The alternative is to continue to provide the subject access we currently provide, whatever that is really achieving, while recognising that subject access to the collection is the area where most clients have most problems, and that this factor is going to continue to frustrate clients and possibly deter them from using libraries.

And just as the major directions of the last several years, outlined in the first section of this paper had a strong emphasis on sharing, the direction I am proposing revolves, for the immediate future, around using what we know we can get through sharing, but tailoring our access for the specific benefits of our institution's clients, using our collection, and within the framework of our institution's mission. A sharing of what is involved in developing these new directions may indeed lead back to shared solutions at some stage further on.

But after all, perhaps it is not that we do not know the answer to effective subject access, perhaps we have just misplaced it. After all in 1775 Samuel Johnson wrote:

> "Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information about it."[4]

The best direction that could be taken with subject access in Australia would be one that would result in our confidently being able to add to this statement and announce to our clients:

> And this Library not only has information on the subjects you are interested in, but also provides ready access to that information.

NOTES

1. ABN subject cataloguing manual. (Canberra: National Library of Australia, 1985)
2. The ASCIS subject heading list, by Ruthi Urbach. Unpublished paper available from ASCIS, 325 Camberwell Road, Camberwell, Victoria, 3124.
3. Subject augmentation in URICA, or Archimedes principle applied? by Alex Burne. Paper presented at the 1985 National Australasian URICA Users' Group Conference, September 18–19, 1985, Perth.
4. The Oxford dictionary of quotations. 3rd Ed. (Oxford: Oxford University Press, 1979) p. 276.

005.BIBL-1.E

RETROSPECTIVE CONVERSION IN NATIONAL AND RESEARCH LIBRARIES: THE AUSTRALIAN EXPERIENCE

Peter Haddad, *Chief Librarian, Control, National Library of Australia, Canberra, A.C.T. 2600, Australia*

ABSTRACT

Automated systems in libraries often require that the records for a large part of the collection be in machine-readable form. In Australia, there has been an upsurge in retrospective conversion activity since 1984, stimulated by a number of factors. Australian national and research libraries have chosen a wide variety of methods to accomplish their retrospective conversion projects. These include in-house approaches, using bibliographic utility services; engaging commercial firms; and various combinations of these. Some problem areas exist, but the Australian scene remains one of vigorous activity. There is scope for the national coordination of retrospective conversion efforts.

Retrospective conversion has been defined as:

> The process of converting to a machine-readable form the

records in a manual or non-machine-readable file that are not converted through day-to-day processing.[1]

Retrospective conversion is a major and costly undertaking for a library, yet in Australia libraries of all sizes and types are increasingly engaging in this activity. Nearly all of the national, state, and larger research libraries have either begun, completed, or are in the midst of some form of retrospective conversion.

The increase in retrospective conversion is a direct result of developments in automation in Australian libraries. Automated circulation systems, inter-library loan systems, on-line public access catalogues, and on-line interaction between library systems for resource sharing, all require large data bases or machine-readable records. To run any of these systems effectively, the core portion of a library's collection must be on-line. Retrospective conversion also provides individual libraries with the opportunity to amalgamate a number of separate bibliographic files into a single, integrated system, with the possibility of upgrading and standardizing cataloguing data. Despite an economic climate of continuing financial restraints, Australian library administrators have given retrospective conversion projects a high priority, seeing them both as necessary prerequisites for the implementation of automated systems, and as investments in the future.

A number of factors have combined to make retrospective conversion easier for Australian libraries. Australian national, state and research libraries were early users of machine-readable cataloguing, and their current catalogues are largely in machine-readable form. National cataloguing distribution systems such as the Australian MARC Record Service (AMRS), established in 1974, distributed MARC records from the Library of Congress, the British Library, and the National Library of Australia, widely throughout Australia. The growth of a national on-line shared cataloguing network, the Australian Bibliographic Network (ABN) has provided a rich source of machine-readable cataloguing from a variety of national agencies, supplemented by original cataloguing input by ABN participants, since 1981. A 1986 survey[2] by CAUL (College and University Librarians) of nineteen Australian university libraries revealed that of a combined 12 million holdings, over 4,600,000, or 38.85%, were in machine-readable form. The proportion is undoubtedly higher today. Alternative and supplementary sources of MARC records such as REMARC have become available to Australian libraries. Overseas bibliographic utilities such as OCLC and commercial bibliographic services such as Blackwell North America have extended their services to the Australian market. Australian commercial firms such as SAZTEC and AMARC Data International Pty Ltd, specializing in handling the retrospective conversion of library data have established themselves. Optical character recognition (OCR) provides a possible new technology for performing retrospective conversions from printed cards.

Most Australian national, state and research libraries regard the ideal situation as one in which all their holdings are in a single machine-readable file, but this is generally beyond the resources of the individual library. Nevertheless most Australian libraries have been engaged in retrospective conversion projects since 1984, and many have converted substantial portions of their collections. The University of New South Wales Library, something of a pioneer in this field, computerised its catalogue in a project beginning in the early 1970s and ending in 1980. This retrospective catalogue was loaded to ABN from 1983-1986. The State Library of Queensland has converted virtually all of the Public Libraries Branch, 85% of the Reference Library's public catalogue, and 75% of the John Oxley Library's catalogue; the State Library Service of Western Australia has planned a staged programme which began in 1984, and is continuing with a number of categories of material like Fiction, Junior and Adult non-fiction completed. Many libraries have chosen parts of the collection for conversion because of automation requirements. The Baillieu Library at the University of Melbourne is converting its current serials holdings in preparation for the installation of an automated serials management system, while the National Library of Australia is also retrospectively converting its serials holdings in advance of the introduction of an integrated library management system. Some retrospective conversion activity is related to the intrinsic value of the collection. For example, the National Library of Australia has retrospectively converted its holdings of Australian items published between 1901 and 1950 contributing these records to the National Bibliographic Database, and producing a printed bibliography in the process. Earlier Australian imprints in the National Library of Australia's catalogue, covering 1788–1850, have similarly been converted.

Australian libraries are faced with a wide variety of methods for undertaking retrospective conversion. These include in-house conversion using staff and equipment specifically hired for the purpose, making use of the data base of a network or bibliographic utility, using the services of a commercial bureau, or a combination of these methods.

Each approach has its own advantages and disadvantages. An in-house approach, for example, using microcomputers or optical disc based systems is costly in terms of finding funds, time taken to set up the project, provision of equipment, hiring and training of staff, but gives a conversion over which the library has a high degree of control, and is likely to provide a retrospective conversion of high quality. Many Australian libraries have used this approach, sometimes in combination with another, to handle data for which a MARC record cannot be readily obtained. For example the State Library Service of Western Australia obtained records from commercial sources, but keyed machine-readable records for the remainder of its main catalogue entries using contract labour. For national, state and research libraries whose files have a large amount of local cataloguing, and whose records are of varying age and quality, this approach allows a high degree of professional intervention to prevent any deficiencies in the data from assuming major proportions in the retrospective conversion.

For most Australian libraries, the services of the national network, ABN, figure largely in any retrospective conversion plans. Libraries which are on-line participants, can always convert in-house using terminals giving on-line access to ABN. Very early in its history, ABN began to offer a batch retrospective conversion service whereby participant files were accepted for conversion either by matching a control number and adding holdings, or by matching author and title, and subject to records conforming to a certain minimum standard of fullness and quality.

The use of a commercial bureau to assist with the retrospective conversion is one of the most common methods used by Australian libraries. It is the cheapest, and sometimes the fastest way, and does not divert as many library staff from the regular operations of the Library. It involves the library staff in liaison with the bureau, sending cataloguing data, and checking the results.

The CAUL survey of nineteen Australian university libraries probably reflects the picture for most Australian libraries. In-house methods were used by 13 libraries, the services of a commercial bureau by 4, and on-line conversion using the national network had also been used by 4 libraries, indicating that a number had used more than one approach.

Difficulties in retrospective conversions generally relate to the data to be converted. Australian libraries have experienced problems in retrospectively converting certain categories of materials such as non-roman script materials. Brief records created for older, or less used portions of the collection have resulted in these portions remaining outside the retrospective conversion process, pending some kind of upgrading. In particular, older catalogues such as that of the National Library of Australia, which dates from the early part of the century to 1967, present some formidable problems of editing and upgrading.

Another area of difficulty is the upgrading of standards for the data retrospectively converted. While descriptive cataloguing standards such as pre-ISBD punctuation may easily remain, authority [illegible]ding upgrading is much more necessary for integration in one [illegible]et the magnitude of the conversion often prevents complete aut[illegible]rity control follow-up. Again, the picture is varied, with some libraries including this as part of their retrospective conversion process, and others only able to upgrade selected types of headings.

The Australian experience of retrospective conversion is one of vigorous and expanding activity. While individual libraries have formulated detailed and well-considered retrospective conversion programmes for their own collections, co-ordination on a national scale has not really been attempted. The CAUL group of College and University Librarians circulate reports and papers which keep members informed of developments in this field, but there has been little discussion of collaboration at the national level. Australian libraries could benefit from a programme such as the Program for Coordinated Retrospective Conversion of Monographs in North American Research Libraries (the ARL Recon Project), which aims to coordinate the retrospective conversion of the remaining 6-7 million records for monographs not yet in machine-readable form. There is scope for major Australian libraries to identify areas of strength in their collections using the RLG Conspectus levels, and to concentrate their efforts in specific areas of subject strength. The organizational process of such an undertaking, while complex, would have the benefit of avoiding duplicate effort. With one network in Australia, ABN, dominant, the potential is there for the distribution and sharing of records quickly.

References

[1] The ALA Glossary of Library and Information Science. American Library Association, Chicago, 1983. p. 194.

[2] Retrospective conversion — Australian University Libraries. Summary of position as at 30 June 1986. CAUL, Melbourne, 1986. (Unpublished paper).

018-BIBL-2-E

RETROSPECTIVE CONVERSION IN THE LIBRARIES OF SOCIALIST COUNTRIES: A SURVEY

N.P. Igumnova, *Director, All-Union State Library of Foreign Literature, Moscow, USSR*

ABSTRACT

The largest libraries of the socialist countries started the realization of cumulation of current machine-readable information and the elaboration of the problem of retrospective conversion. On the regional level this activity is coordinated by the special Working group of experts in the field of electronic data bases.

The libraries of Bulgaria, DDR, Hungary, Czechoslovakia, Poland and the U.S.S.R. create the electronic data bases according to the following principles: national and international compatibility, usage of international standard number and IFLA regulations on UNIMARC format and on bibliographical description of printed matter. The electronic data bases are formed of the basis of utilization of large EC computers, mini- and personal computers.

IFLA Section on Bibliography, Division of Bibliographic Control, defined the conversion of large library catalogues, current, retrospective and thematic bibliography in machine-readable form as one of the perspective directions of its activities.

The socialist countries pay much attention to the problem of cumulation of current bibliography and retrospective conversion.

The special Working Group on library technology, including the most qualified specialists, has been formed under the auspices of the Ministries of Culture of the socialist countries. The aims of the Working Group stipulate the elaboration of the problem of computerization in the research libraries. Other professional groups, e.g. experts on bibliography, cataloguing, etc., also dwell upon the problem of conversion of national bibliography.

At the 8th meeting of experts of the socialist countries on the problems of national bibliography (Leipzig, 1987) the experts from Bulgaria, DDR, Hungary, Poland, the U.S.S.R. discussed the results of the formation of automated data bases.

This survey of the present state in the field of the formation of machine-readable bibliographic indexes and retrospective conversion of catalogues in the socialist countries is made on the basis of information received from the national libraries of Bulgaria, DDR, Poland, Czechoslovakia.

In Bulgaria the necessary prerequisites are created for the availability of information on Bulgarian literature and for the development of international exchange of machine-readable bibliographic data.

The Cyril and Methodius National Library has formed the system of retrospective bibliography. It includes: books — since 1806, thesis — since 1929, periodicals — since 1844, articles — since 1806 and bibliographical publications — since 1852, all in all 180 thousand bibliographical items. In the framework of this system the retrospective bibliographical indexes are published. But the retrospective conversion of catalogues and bibliographies will be fulfilled only after the machine-readable data bases of current national bibliography are created. This programme, planned for 1988–1990, provides for the conversion of retrospective, cumulative and union indexes in machine-readable form on the basis of current national bibliography data, accumulating in machine-readable data bases.

The bibliographic description of all the publications for the machine-readable data bases is made on the basis of national standards in accordance with ISBD recommendations. For all the publications, except periodicals, the exchange format is UNIMARC. The classification scheme of current national bibliography will be used in the process of retrospective conversion during the bibliographical data input. The information on national book publications will be provided on-line as well as using diskets, disks, microfiches. The Bulgarian specialists believe that efficient on-line services could be established only after the automated network with direct remote access to data bases is created in the country. The computation centre of the Cyril and Methodius National Library acquires new technology that is necessary for the formation of machine-readable data bases.

This process began in Hungary in 1977, when the automated system of Hungarian national bibliography was put into operation. The retrospective cumulations are formed on the basis of current information. In the libraries of Hungary the current and retrospective information banks are created, on the basis of which they prepare bibliographical indexes and printed catalogues. The service of selective information distribution is established. The National Szechenyi Library plans to form the machine-readable retrospective union catalogue of printed publications that supposes to have information output on micro-carriers. In 2–3 years this library will have a large computation centre using computers with powerful memory. That will help to collect retrospective files of Hungarian and foreign books and serials and to establish on-line services.

It was mentioned in the press that the Hungarian Central Technical Library now carries out the retrospective conversion of the union catalogue of foreign technical journals and also of the foreign books file. Other libraries (Library of the Central Statistical Office, Library of the Szeged University) fulfill retrospective conversion in the appropriate fields. The formation of machine-readable data bases in Hungary is performed according to ISSN, ISBN/the numbers are centrally confered in the country), subject headings, authors, titles, UDC.

In the libraries of Poland machine-readable data banks are formed on the basis of current bibliography. The National Library in Warsaw has worked out the programme of retrospective conversion of catalogues and retrospective bibliography. The necessity of forming a modern catalogue accounts for the decision to convert the name catalogue of serials of the National Library (started in 1985) into machine-readable form. Today this catalogue includes 45 thousand titles of Polish and foreign serials, published after 1800. The annual increase of information — 10 thousand entries. The format employed is SABINA. The structure of the bibliographical entry corresponds to the ISO 2709 standard (MARC format). The entry is worked out on the basis of UNIMARC. The subject indexer is used as information-retrieval language. About 2 thousand entries are registered now on flexy disks; it includes 13 thousand annual sets of issues. The retrospective conversion of the union Polish serials catalogue was started in 1986. The catalogue includes about 40 thousand titles, published in Poland and also in Polish language all over the world.

The machine-readable retrospective data bases of Polish serials is formed for the purpose of preservation and extensive access to information. It is planned to establish the on-line service and the distribution of data bases on modern information carriers.

The system of information input is as follows. The POLARKA format is used, specially created for this catalogue. The bibliographical data, previously put in POLARKA, is to be automatically converted in micro-computer AMSTRAD. The subject indexer is used in information retrieval. ISBN, state library index, state bibliographical description standard on printed matter are used in automated system of the National Library of Poland.

The activity on the formation of the machine-readable data banks of current national bibliography and on retrospective conversion of accumulated information file is carried out in the libraries of DDR. Now 160 data banks are used, 120 exist as catalogues and subject files retrospectively converted in machine-readable form. The Deutscher Staatsbibliothek fulfils the retrospective conversion of the union catalogue of foreign periodicals and of the union DDR newspapers catalogue.

The Deutsche Bücherei in Leipzig — the centre for national bibliography — converts the cumulated volumes of national bibliography on magnetic disks. In 1971 this library began to produce the bibliographical indexes using new technology, according to which the retrospective bibliography "German books index (1981–1985)" is prepared. In printed form it will consist of 7 volumes of general text and 4 volumes of indexes. The work on the machine-readable "Academic Publications Annual" is in progress. The Zentralinstitut für Bibliothekswesen and the Zentralinstitut der Information and Dokumentation der DDR render scientific and methodological assistance to German libraries in the field of the formation of machine-readable data bases and retrospective conversion.

Other institutions fulfil retrospective conversion of catalogues on different branches of science and form union data banks that are extensively available to specialists. For example, the Bibliothek des Ökonomischen Bruno Zeger — Instituts works on the retrospective conversion of catalogue on economics, the Bibliothek der Tech-

nischen Hochschule in Leipzig — on building, the Bibliothek der Technischen Hochschule in Merseburg — on chemistry.

The present information policy development in CSSR provides for the creation of the union information collection system originated from the machine-readable data bases on the basis of national bibliography.

The following main principle is realized: compatibility of automated information data bases on the national and international level.

The integrated automated library information system is now being worked out in Slovakia. It provides for the creation of current Slovak national bibliography data base and for the automatization of retrospective national bibliography. This research task is realized in Slovak National Library. The main aim is the ensuring of unity in library and bibliographical processing of documents in all the libraries of Slovakia. The Slovak national bibliography will act as the central data base of the integrated system which is to be put in progress after 1990. It is much more difficult now to solve the problem of retrospective conversion of catalogues and bibliography in the research libraries of Slovakia. With this purpose the experts from CSSR update the state standard on cataloguing and bibliographical description on the basis of ISBD. The work is to be completed by 1990. In a number of libraries the specialists have already started to work on the retrospective conversion of catalogues: the Slovak Technical Library has realized the retrospective conversion of the Standards catalogue and the journals catalogue. Unlike the traditional description, the retrospective conversion of the standards catalogue has the extended description structure and the variable length. The retrospective conversion of this catalogue represent the data file, formed in accordance to numbers, subject headings and systematized on branches of science. The standards file on librarianship and related subjects is alloted. This data base now enumerate more than 200 thousand descriptions.

The record is carried out on flexy disks Consul 2712 and Consul 2713. The data is converted on magnetic tape for EC 1040 computer.

The central data base of the State Library of the Czech Socialist Republic act as the basis of the national information library network. The preparation of the Czech book bibliography by means of computer technology started in 1983. The process includes phototypesetting of the bibliographic catalogue and the indexes. The data base of Czech books enumerates 30 thousands entries. As an experiment it is planned to use this file in on-line retrieval. For this purpose personal computers are used. The bank of equipment — Mini-micro version CDS/ISIS was purchased in 1986 from UNESCO. On the basis of this technological equipment the inverted data file is formed with the purpose of information retrieval according to selected elements, indexes. The linguistic provision is as follows: subject headings language, key words and name indexes, UDC. These elements are combined with the help of logical operators. In present (since 1983) the machine-readable data base on culture is formed. Experimentally the machine-readable data base of Czech journal articles must be put in operation in 1990. It is planned to create the data base on thesis.

In CSR the compatibility and interconnection of national and regional bibliography data bases are realized. In the purpose of international exchange of bibliographical information the problem of compatibility of entries is now being solved on the basis of inculcation the international standards. In addition to this activity the national data is converted in the international machine-readable data bases. It is important to mention the work being done in the field of author description and subject processing of documents.

The entries are transmitted to central data base of the International system on Social Sciences of the Socialist Countries.

In the USSR in the framework of the State Automated Scientific Information System the formation of the machine-readable data bases on different branches of science started after 1970. Files and data bases of current domestic and foreign documentation are created on the basis of information accumulation over 5 & 10 years. These collections are gradually converted in retrospective collections and act as on-line information retrieval systems with different information volume available. Within this programme in 1986 the preparation for the operation of the automated library information system (libraries of the USSR Ministry of Culture) was started. The technical feasibility of this programme is now being discussed. This multifunctional system will provide for the creation of all-union data banks on the basis of the largest libraries. It is planned that the V.I. Lenin State Library of the U.S.S.R. will exercise the functions of the coordinator of this system, for it has the largest collections and catalogues of the domestic and foreign literature in the country. The machine-readable data bank of foreign fiction and scientific literature on the humanities will be formed in the All-Union State Library of Foreign Literature. Regional machine-readable catalogues will be organized in the state libraries of the republics and regions of the country. Due to this programme it is planned to fulfil the retrospective conversion of large library catalogues. It will help to use the information resources of the country. The technological equipment provides for the direct access to the external storage and for the information output on paper and micro-carriers as well as the duplication, using the effective means of reprography and small-graphic-arts industry.

The total information volume on the machine-readable carriers in the process of retrospective conversion of catalogue is defined at the rate of 20 years of information storage. The linguistic complex includes: bibliographic description language, classification information-retrieval language, subject headings and descriptors language, figure codes, corresponding tables of library classification and state rubricator, preprocessing and exchange formats. The software includes such means as applied programme package CDS/ISIS/EC mini-& micro and additional programme modules.

The V.I. Lenin State Library works at the retrospective conversion of the repertoire of the XIX century Russian books. The electronic union catalogue of the XIX century Russian books will be formed. The "Union catalogue of the XIX century Russian books (1801–1825)" will include 93–94% of books of the above mentioned period. The bibliographic description is fulfilled in accordance with the "State standard GOST 7.1-84 on bibliographical description of documents & the regulations on bibliographical description of old books and the international IFLA regulations on bibliographical description of old books. About 30 bibliographical aspects are defined: contents, design, printing features, single copy features, etc. The complete control and precise bibliographical description provide for the creation of the proper base for the historical, cultural and bibliographical research. The preservation of this precious information on magnetic carriers gives the opportunity to include it in the machine-readable catalogue of the all-union collection of book monuments.

The utilization of large EC computers, micro- and personal computers produced in the U.S.S.R., Bulgaria, DDR for the creation of bibliographical machine-readable data bases is the characteristic feature of this process in all the socialist countries. The use of computers in Bulgaria, DDR, Poland, Hungary, the U.S.S.R. and Czechoslovakia made it possible to start the formation of machine-readable data bases of current national bibliography. The project programmes on the creation and exchange of machine-readable data bases in the national libraries of the socialist countries ensure the development of automated current bibliography as well as the process of retrospective conversion of large library catalogues. By all means it will greatly increase the availability, effective use and preservation of information and stimulate the international exchange of bibliographical data.

Acknowledgement

The author of this paper highly appreciates the assistance rendered in preparation of the paper and also the information on the problems of retrospective conversion in the libraries of the socialist countries which was received from: K. Zotova (National Library, Bulgaria), B. Karamats (National Library, Poland), d-r M. Kovachka & M. Bielik (Slovak National Library, CSSR), d-r H. Pomassl and H. Rotcsch (Deutsche Bucherei, DDR), A. Vostrikov I. Polonskaja & P. Berdnikov (V.I. Lenin State Library, U.S.S.R.).

References

[1] VIII Rabochaja gruppa Ministerstv cultury sozialisticheskich stran. Informazionni bulletin. — N 2. Bibliotechnaja tehnika i tehnologija. — Martin: Matica Slovenska, 1986. — 140 c.

[2] VIII meeting of experts of the socialist countries on the problems of national bibliography. — Leipzig, 29.09.-3.10.87: Papers by: Zonneveig P., Zotova K., Igumnova N., Kovachka M., Pomassl H., Radova I., Ramlau-Krehovska K., Rost H.

Bulgaria

[3] Atanasova D. Avtomatizirani bibliotechno-informazionni tehnologii i prilozenietoim v CB na BAN //Bibliotekar. — Sofia, 1987. — God. 34, N 6. — C.2-5.

Hungary

[4] Nyáriné-Grófcsik E., Szucs E. A Magyar Nemzeti Bibliográfia számítógépes rendszere //Könyvtari figyelo. — Budapest, 1982. — 6.sz. — 573–590 old.

DDR

[5] Honeit W. Das Organisation — und Rechzentrum des Deutsche Bücherei //Zentralblatt für Bibliothekswesen. — Leipzig, 1987. — Jg. 101, N 9. — S.404–406.

[6] Penzold L., Kittel P., Wehmeyer A. Die Verfilmung des Alphabetischen Hauptkatalogs der Deutschen Staatsbibliothek Arbeitsvorbereitung, Technische Durchfuhzüng, Möglichkeiten der rechnergestützten Recherche //Zentralblati für Bibliothekswesen. — Leipzig, 1986. — Jg. 100, N 9. — S.391–400.

[7] Plesske H.-M. Einige Bemerkungen zum EDV-Projekt Bibliographieherstellung der Deutschen Bücherei //Fontes artis musicae. — Basel; London, 1982. — Jg. 29, N 1/2. — S.51–54.

[8] Rost G., Hahn F. Ansätze zu einer ökonomischen Bewertung bibliographischer Tätigkeit //Bibliographie aktuell. — Leipzig, 1982. — N 2/3. — S.77–88.

[9] Rost G. Zur Terminologie im EDV-Projekt Bibliographieherstellung der Deutschen Bücherei //Bibliographie aktuell. — Leipzig, 1982. — N 2/3. — S.7–8.

[10] Schultze S., Meissner A., Kahnt G. Die rechentechnische Entwicklung des EDV-Projekts Bibliographieherstellung //Bibliographie aktuell. — Leipzig, 1982. — N 2/3. — S.9–34.

Poland

[11] Avtomatizirovannaja informazionnaja sistema Nazionalnoj biblioteki v Warszawe /sost. I.F. Tutina; Nauchn. red. S.V. Pushkova. — M., 1987. — 7 c. — (Cultura i iskusstvo za rubezom. Ser.: Bibliotechnoje delo i bibliographija. Ekspress-inform./GBL, Vyp.3).

[12] Wilgar J. Zaklad Bibliografii Polskiej 1901–1939. dalsze losy edytorskie Bibliografii Polskiej 1901-1939 //Biuletyn informacyjny Biblioteki Narodowej. — Warszawa, 1985. — N 3/4. — s.12–14.

U.S.S.R.

[13] Alekseev N.G., Vostrikov A.N. Sostojanie i osnovnye napravlenija raboty po avtomatizazii Gosydarstvennoj biblioteki S.S.S.R. im V.I. Lenina //Tr. /Gos. biblioteka S.S.S.R. im V.I. Lenina. — T.21: Problemy avtomatizazii i mehanizazii bibliotechnoj raboty. — M., 1985. — C.92–105.

[14] Igumnova N.P., Majo-Znak E.O., Nazmutdinov I.K. Mezdunarodnoje sovescanije ekspertov sozialisticheskih stran po voprosam nazionalnoj bibliographii /Moskva, 9-11 junja 1981 g.) //Bibliotekovedenije i bibliographija za rubezom. — 1982. — N 88. — C.81–92.

[15] Petrenko G.A., Trushkina Z.N. Osobennosti obrabotki izdanij v AS GBL //Tr. /Gos. biblioteka S.S.S.R. im V.I. Lenina. T.21: Problemi avtomatizazii i mehanizazii bibliotechnoj raboty. — M., 1985. — C.106-109.

Czechoslovakia

[16] Ďurovčík Š. Nadvaznosť spracovania národnej retrospektivnej bibliografie na výstopy z AIS SNB //Knižnice a vedecké informácie. — Martin, 1984. — Roč. 16, N 6. — S.258-263. — Bibliogr.: s. 263 (11 ref.).

[17] Kovačka M. Ku koncepčným otázkam výberového radu AS SNB // Knižnice a vedecke informacie. — Martin, 1984. — Roč. 16, N 2. — S.78–82.

[18] Niektoré súvislosti výstavby bázy dát AS SNB s výstupmi výberového radu slovenskej národney bibliografie //Knižnice a vedecke informacie. — Martin, 1984. — Roč. 16, N 3. — s.111–114. — Bibliogr. : s.114 (6 ref.).

43-BIBL-3A-E

PROBLEMS OF RETROSPECTIVE CONVERSION IN NATIONAL AND RESEARCH LIBRARIES: THE EXPERIENCE OF THE WESTERN COUNTRIES

— The conversion of the British Library Catalogue to machine-readable form

Pat Oddy, *Head, BLC Conversion Project, British Library Bibliographic Services, London, UK*

ABSTRACT

British, French and Spanish experiences in retrospective conversion well illustrate difficulties encountered by European countries in this field. Old catalogues of each of these countries have a peculiar structure and require a specific treatment. Due these circumstances, each country solves problems in its own way, and as compatibility of the getting database must be ensured, technical recommendations have to be rapidly established.

INTRODUCTION

In countries where edition is an ancient tradition, as European countries, recording old material in machine-readable database accessible to all is regarded as prime importance. The basic principle of a "retrospective conversion" is to use the existing traditional catalogues as a basis for conversion into machine-readable records, without books in hand recataloguing. Structure of original catalogues takes a predominant part in retrospective conversion. Many European countries have ancient catalogues which have been elaborated at a time when one could not even conceive their computerization: each catalogue has its internal logic and sets specific problems at the time of its conversion. So great is the cost that the adopted solutions are not always the best ones, but the cheapest. Then countries which have here and now undertaken retrospective conversion of their catalogues, have searched for appropriated solutions to their local problems: the account of the British, French and Spanish experiences in this field is an illustration of that. Divergent national solutions could be a disadvantage for transboarder data flow; then it is not surprising that this problem is at the heart of several meetings at European scale since 1987. So, after Western European national libraries directors' Lisbon meeting about retrospective cataloguing, it is now the European Council which has created a working group on this matter. Therefore study of our Section contributes to the reflexion.

(original text in English)

In January 1987 the British Library entered into a contract which provides for the conversion of the 360 volumes of the British Library general catalogue of printed books to 1975 (BLC) to machine-readable form. The conversion will be complete in 1991, when the major catalogue of the Library, which has existed for the general public in published book form since the beginning of the century, will be available as an on-line database accessible not just to readers in the new British Library currently under construction in central London, but also to remote users through the Library's own automated information service.

Although often referred to as a retrospective conversion, this project differs from most such conversions for two major reasons. Firstly, because it is a move from a print, rather than a card based catalogue, and secondly because of the nature of the collection which that printed catalogue represents. In common with most national libraries, the development of the collection over time means that it is heavily weighted towards older materials. The practice of matching and down-loading existing records from other databases consequently produces low returns and the Library has had to face the task of a unique conversion project if it is to reap the benefits of the speedy and comprehensive access to its collection which an automated catalogue will bring.

At the heart of this conversion project is the data conversion specification, produced within the British Library, which details the way in which the printed catalogue entries are to become machine-readable unit records. To appreciate the problems which the specification has had to address, a brief description of the printed catalogue is necessary. BLC consists of about 4 million main book entries, which contain the data elements such as title, imprint and shelfmark by which a work and its various editions and copies can be identified and located in the British Library collection. Imposing a framework on these book entries are the structured elements of the catalogue: the headings and subheadings which determine both the alphabetical linear sequence of the printed catalogue, and the subsequences in which the entries related to particular complex headings, such as prolific authors or government entities, may occur. The catalogue also contains finding references and information entries which guide the user to authoritative forms of heading and explain the sequence of entries under complex headings. Headings appear once in the catalogue and are presumed to apply to all the following book entries until another heading appears. Subheadings follow this pattern as well, with the consequence that the full hierarchy of a heading and its subheadings may need to be picked up over a wide range of pages in BLC. It is almost impossible for a two-dimensional, printed sequence to express the relationships between these isolated headings and subheadings clearly, and one of the major problems for the conversion project has been the need to bring together these elements to form hierarchically logical headings and uniform titles. A final and important element of the printed source catalogue is the existence of cross-references,

which provide access to the main book entries through the headings for related persons or corporate bodies.

From this description it will be clear that the current concept of a unit record, as exemplified by a catalogue card with tracings or a MARC record, is absent from the printed BLC. It will be the final element in the conversion process to carry out what we are calling the compression of the converted file: this means that the cross-references will be matched by shelfmark with the main book entry to which they refer, and the heading and any subheadings from the cross-reference transferred to the main entry as added entry access points. Only when this final matching process is at an end will the conversion process be complete.

Data capture and conversion of BLC is being carried out by Saztec Europe, working from their facility in Ardrossan, Scotland. Using the data conversion specification, they first identify the type and structure of each entry within a volume and then define the data contained in each entry into basic bibliographic elements such as title, imprint and collation, and series. Headings and subheadings are captured and sent as a separate diagnostic output to the BLC Conversion Project team based in BL Bibliographic Services. There the individual subheadings are allocated a letter or number code which both defines them as either structured title or corporate body subheadings and ensures their correct position in the sequence of subheadings building a heading or uniform title. The marked-up list is returned to Saztec and the data merged with that for the descriptive elements of the volume.

The data is then manipulated in such a way as to produce unit records which are coded to conform broadly with the UKMARC exchange record format. Obviously the records will not comply with the Anglo-American Cataloguing Rules which the UKMARC format was designed to hold — the original catalogue entries were created using British Museum cataloguing rules — but they will allow a level of standard MARC processing.

Finally, the records are output to magnetic tape, each tape holding data for one volume of the printed catalogue. Within the BLC Conversion Project team, quality control operations are carried out on a selected sample of the records to ensure that the required levels of accuracy have been achieved.

The BLC conversion project addresses a problem faced by many national and academic libraries, particularly perhaps those in Europe: the need to create functional records for a machine-held database from a catalogue with an internal structure designed for a print-based, linear medium. The results of the project to date show that investment in developing a serviceable conversion specification can pay remarkable dividends in opening up a printed catalogue to multiple search strategies.

43-BIBL-36-E

PROBLEMS OF RETROSPECTIVE CONVERSION IN NATIONAL AND RESEARCH LIBRARIES: THE EXPERIENCE OF THE WESTERN COUNTRIES
— Integrating the Catalogue général des livres imprimés 1970–1979 of the Bibliothèque nationale into the BN-OPALE database

Annick Bernard, *Chef de projet, Bibliotheque nationale, Paris*

(original text in French)

At the beginning of 1982 the Bibliothèque nationale completed its Schéma directeur Informatique or computerisation plan, which brought forward the need for setting up a large bibliographical database. This was to be the BN-OPALE data-base.

At the same time, the management had to decide by which method the Catalogue général 1970–1979, consisting of a file of 1 million cards, corresponding to 450,000 works, should be published.

A straightforward photographic reproduction of the cards, which was the procedure used for the 1960–1969 catalogue, would have resulted this time in a catalogue of some 50 volumes at a very high cost price, so this method was quickly rejected.

It was decided that the entries should be keyboarded so that they could be transferred later to the BN-OPALE data-base. The catalogue would emerge as a by-product of this keyboarding operation.

The problem was not an easy one:

— the 1970–1979 card file had been organized so that it could be reproduced in the same way as the previous catalogue, with cards arranged in alphabetical order for each entry in a record, and without repeating the headings on the main entry.

Now, however, in order to transfer the entries to the data-base, we had to reconstruct a single record, together with its added entries.

— The Bibliothèque nationale only had a small team of at most 13 people at its disposal, which had to concentrate mainly on scholarly revision: a large card file representing 10 years' cataloguing inevitably required standardization of headings and cataloguing practice. There were no facilities for the actual preparation of the entries (identifying the various types of cards, coding for keyboarding) to be carried out at the Bibliothèque nationale.

— a further constraint was that only an access code needed to be keyboarded for records published in the Bibliographie de la France between 1975 and 1979, allowing retrieval of each entry in the existing computerised file and its incorporation into the catalogue.

The employment of a contractor specializing in record distribution was briefly considered as part of a composite solution, but this was rejected as the overheads would have remained unacceptably high.

The agency that we retained at the end of 1984 to complete the work in 3 years had the following remit:

— preparation of the card file for keyboarding
— duplicate keying
— computer programmes and information processing
— conversion to INTERMARC format
— provision of an editing file

Preparation for keyboarding

As a preliminary, we had to make a detailed analysis of the file and define a precise set of rules.

The agency staff therefore learned to differentiate between the various types of cards and headings, to pick out the title proper by its typography, etc . . . and to take account of changes in cataloguing rules that occurred in 1975 with the introduction of ISBD, which mean that the same information is not always shown in the same way.

Bibliothèque nationale staff only had to mark doubtful cases that might lead to errors being made.

Duplicate keying

Thanks to duplicate keying, few keyboarding errors arose. Bibliotheque nationale staff proofread the keyboarding exhaustively whenever they had the time. Otherwise, they confined themselves to dealing with errors messages produced by automatic checks and to a rapid check of essential elements.

Programmes and information processing

In the absence of an LC-style unique record control number, related entries are linked to the main entry by means of the pressmark, which, in theory, identifies the record unambiguously.

At the end of the process aimed at achieving this linkage, proofsheets and lists of queries are produced, which are carefully checked by the Bibliothèque nationale staff, who make any necessary corrections on worksheets.

Conversion to MARC format

Entries prepared for keyboarding are roughly structured with the aid of INTERMARC tags. During conversion this structure is refined by the addition of sub-field codes and the automatic resetting of certain data.

When it is loaded on to the BN-OPALE data-base further improvements will be added by profile.

Provision of an editing file

The editing file provided by the service agency will show the 1970–1979 catalogue as it was in card-index form.

The publication, to appear at the end of 1988, will not be on paper, but on COM microfiche. It will thus complement the microfiche edition of the Catalogue général published in 1987.

Assessment of the project

At the time of writing we can say that the project has been a success. The keyboarding has been completed and the successive stages of data processing and subsequent checking will be finished in the Summer of 1988. The transfer to BN-OPALE will be effected between Summer 1988 and the beginning of 1989.

At a cost of 8.7 million francs, i.e. 8.70 francs per original card entry, the Bibliothèque nationale will have increased its bibliographical data-base by 350,000 records, which it will be able to put at the disposal, not only of its readers, but also of other libraries.[2]

Such a large-scale transfer of more than 700,000 author entries, plus about 100,000 cross-references, is bound to have an effect on the status of the personal and corporate name authority files. The staff working on the 1970–1979 catalogue have been liaising closely from the beginning with the managers in charge of the authority files in order to preserve their coherence as far as possible.

The state of the original card file and the desired end-product dictated the choice of methodology.

Pre-requisites for success

— the will to finish within a reasonable period of time
— the necessary financial resources
— a soundly prepared project
— ample data processing facilities
— a motivated staff willing to accept a constraining work schedule (40,000 records per month to be handed over to the service agency) and a stable work force on both sides (Bibliotheque nationale and external staff).

Other projects

The Bibliothèque nationale, which is at present setting up a new automation plan covering the next five years, will have to decide on the viability of recording and converting to INTERMARC format the 255 volumes of its Catalogue général des livres imprimes and its card supplements, i.e. 3.5 million entries.

As far as the anonymous antiquarian (16th–18th c.) entries are concerned, the re-working of the existing summary card catalogue was undertaken in 1984: the entries are being re-catalogued from the original books — so it is a long-term project — and keyboarded by a service agency. The entries will be loaded on to BN-OPALE. A microfiche catalogue, with successive updates, is being published annually.

Bibliography

A. BERNARD and M. CHOUDEY, L'opération N.C.G. ou l'intégration du Catalogue général 1970–1979 dans la base BN-OPALE. In: Bulletin d'informations de l'A.B.F. (Association des bibliothécaires français), number 137, 4th trimestre 1987.

Notes

1. In 1988 the BN-OPALE data-base contains almost 600,000 bibliographical records, comprising all French books and serials since 1975 and foreign books and official publications since 1983. In 1989 it will reach 1 million records.
2. The BN-OPALE records are extracted and provided on magnetic tape on request. The current output is available on subscription. Soon a CD-ROM will allow an extensive distribution of the records in the Bibliographie de la France from 1970.

43-BIBL-3C-E

PROBLEMS OF RETROSPECTIVE CONVERSION IN NATIONAL AND RESEARCH LIBRARIES: THE EXPERIENCE OF THE WESTERN COUNTRIES
— The Spanish experience

Xavier Agenjo, *Chief of the Service of the Union Catalog of Spanish Heritage, National Library, Madrid*

(original text in English)

The process of automation the cataloguing of libraries presents two different problems: that of the automation of cataloguing the material of new accession and that of cataloguing the already existing. In many countries the first one is going to be solved and that is why they are facing the second; however, that is not the case of Spain, and because of it the retroconversion is not a main priority.

The situation is as follows: among the Spanish research libraries, not all the university libraries are automated. As a matter of fact the biggest one, the Library of the Universidad Complutense de Madrid, has not started the process, although it seems to be immediate. In general the bibliographical description of monographies is handled with two different softwares: the DOBIS/LIBIS of IBM and the SABINI, created by the Spanish firm OCSA: both use the MARC format. The Public and University Library of Barcelona, which has its own ad-hoc software, have a different situation.

In a very small scale, all these libraries are recataloguing the existing material, without using the method of keyboarding the existing files, not mentioning the Optical Character Recognition, etc. In no case the use of Databases is contemplated, except in what concerns the Public and University Library of Barcelona, which is in touch with the OCLC. Anyway for the moment the results are not totaly satisfactory, because many records are not found in this database, specially Spanish books records. Bureau services are not considered. Some libraries are thinking in the possibility employing librarianship students to start the process, such as the University Library of Granada, which has DOBIS/LIBIS, and has planned to employ students of its Escuela Universitaria de Biblioteconomía y Documentación to make the bibliographical description of the items in circulation, reserving the pre-1801 for the library staff.

A completely different case would be that of the Consejo Superior de Investigaciones Cientificas (CSIC), a state organization that reunites the majority of the Research Public Centers, which has foreseen, in the frame of its libraries Automation Program, a section dedicated to the retroconversion. It is necessary to point out that the CSIC disposes of the ALEPH program, nevertheless it uses the Library of Congress tapes to integrate records. On the other hand it has a system of sharing retrospective cataloguing, because several of its libraries of Madrid are working already in net. But it is necessary to say that for the pre-1801 prints a different methodology has been followed, cooperating in this case with the National Library.

The National Library has the intention of beginning in 1989 its retroconversion. With that aim, several tasks have to be started: one of these is to implement in the memory of its integrated system SABINA, a development for the National Library of SABINI, the logical records used to elaborate and publish Bibliografía Española, which since 1976 is codified with the IBERMARC format, which derives from the USMARC. The same project exists with the tapes of Bibliografía that from 1969 to 1976, were codified with an ad-hoc format. After that it will be necessary in both cases to add local data.

Another two projects are based in taking advantage of the Spanish Union Catalog of Serials and the Union Catalog of the Spanish Heritage, which is trying to reunite manuscripts, special materials and foreign prints, and Spanish books printed until 1958, date of birth of the Copyright Deposit.

The Spanish Union Catalog of Serials has counted with the ISDS tapes and for the same process as Bibliografía Española, that means, adding local data, will help with the conversions of serials, although the problems of holdings seem to be very serious.

The Union Catalog of the Spanish Heritage has a different methodology: using the ISBD(A), (M), (PM), (NBM), etc, and some adaptations of the USMARC format for manuscripts, special materials and prints, even old prints, a base-catalog has been collected. This catalog includes the stocks of several different libraries, very rich in those materials; among these obviously the National Library. As the bibliographical descriptions follow the standard ISO-2709 it is very easy to capturate them from automated libraries. The Library of the Universidad Complutense de Madrid, which, as I have said below, is not yet automated, it is one of the libraries choosed to become part of the base-catalog. It will be possible next year to take records from the Union Catalog and integrate those records in its automated system. The same goes for the libraries of the CSIC, that are also in the base-catalog.

The Union Catalog is a project of the Ministry of Culture in collaboration with the Autonomic Communities: Spain is politically divided into 17 Autonomic Communities. The Department of the Bibliographical Heritage of the National Library is preparing, as I have explained below, a base-catalog which is sent to the different Cataloguing Centers of the Autonomic Communities with the aim of getting any local data of the already existing-in-the-catalog editions and to add more descriptions of other editions. These new descriptions could be used by the rest of the Communities, libraries, and so on, in an interactive way. As all the process is automated, the retroconversion of that kind of material collected by the Union Catalog can be very fast. That is why it is very important to follow the standards of IFLA, ISO, etc, to avoid the programs of conversion, when possible.

Parallel to this effort the National Library plans to send out immediately its authority files, in Marc format, to any library starting its process of automation, with the intention of making easier, getting homogeneous entries and, connected with retroconversion, allow the search of bibliographical records from its database.

Finally: although the retrospective conversion is not yet a priori-tary task in Spain, it must be said that it is regarded in almost every project of library automation, trying to avoid mistakes that could make it more difficult to future, and taking advantage of every choice.

CONCLUSION

British, French and Spanish realizations in the field of retrospective conversion well prove that adopted solutions are different from country to country, and even from catalogue to catalogue in the same country. Each country, depending on structure of its original catalogue and on funding at its disposal to bring the project to a successful issue, is attempted to use individual rules for conversion. Emerging of non-compatible automated catalogues is a real risk. The number of retrospective conversions in process is a positive point, but to avoid uncompatibility which considerably restrict the accessibility of these catalogues, it seems urgent to elaborate technical recommendations in order to teach countries starting this operation. That is the program the working group on retrospective cataloguing, previously presented in the introduction, set in the European area.

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991